A generic subject: The interplay of morphosyntax and the human conceptual system

Kim K. Fuellenbach

St. Hugh's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2020

Acknowledgements

While it might seem obvious to start your acknowledgments by thanking your supervisor, Matt Husband wholeheartedly deserves my gratitude for being an outstanding advisor. You have shown me endless patience, and I am grateful for your guidance through this process, and of course your excellent feedback over the past five years. Thank you also for your trust in my judgements, be it for experimental design or any of the other innumerable decisions that happen during a DPhil.

I would also like to express my thanks to Aditi Lahiri, who has provided advice for all matters thesis and life-related. I very much appreciate having had the opportunity to work with you and have been a part of the LABlab. I especially thank Aditi and Jenny Culbertson as my examiners for their time and effort to provide me with the input that enabled me to make my thesis the version that it now is.

I am incredibly grateful to Susan Gelman and the Conceptual Development Lab at the University of Michigan. Your support and your hosting me in your lab was an invaluable addition to what I have learned as a linguist with a deep appreciation for the value of psychology studies. Susan, you have in many ways become an unofficial mentor and advisor. Thank you for all your well-formulated feedback, believing in the value of a linguist coming into your lab, and helping me in so many ways to carry out research that seeks to compliment in a small way the impact you have made in developmental psychology. Nicole Cuneo, you are a superstar for keeping on top of everything: from endless stimuli recording to practically setting up studies, and always asking the right questions with an eye for detail. I know you will be incredibly successful in your academic career.

My DPhil and this collaboration would not have been possible without the financial support from the ESRC Studentship for my MPhil and DPhil, the Grand Union DTP Overseas Institutional Visit Grant, the Santander Academic Travel Award, and the Basant Kumar and Sarala Birla Graduate Studentship. Thank you for allowing me to make this DPhil the project that I wanted it to be. Equally, the experimental aspects of this thesis would not have been possible without the discussions in the weekly lab meetings, and the help from a large number of Master's students in the Faculty over the years: thank you, and in particular to Daniel Cornford, Jonathan Knight, and Nina Ubaldi for help with translation and coding.

The last few years have seen me spend most of my time in the lab. In particular to my office mates, desk predecessors, and lunch time co-banterers, thank you.

Discussions with you indubitably made me a more well-rounded linguist. Emily Darley, Zoë Fletcher, Yoolim Kim, Arkadiusz Kwapiszewski, Swetlana Schuster, Ana Werkmann-Horvat, Rebecca Zeng, and Adam Zhang, you've all inspired me with your stories, successes, support, and willingness to go to all the veggie pubs in the evenings and Gloucester Green market over lunch.

My home was without a doubt Worcester: Heather was my very first friend there and it is only fitting that we met in the library. You've provided me with daily support in-person and remotely, tea chats, yoga sessions, book club meetings, and lunch breaks. Hannah, you've not only peer-supported me as a friend but crucially also by letting me use a certain SCR's espresso machine to add to scrumptious Taylor's breakfasts when the goings got tough. Other times, you've introduced me to bread sauce and welcomed me to your home on numerous festive occasions. Thank you for making it my own home away from home. Katie Ball and Lizzie Freeman were the best housemates and always up for a brunch or mid-afternoon cinnamon bun - I miss sharing a kitchen with you. The Worcester College D(r)ea(n/m) Team. It is obviously hard to find a group that would be more qualified to provide ongoing support of any type but you all have outdone yourselves. Here's to my favourite team of strong, independent, courageous, and supportive women: Jo, Shazeaa, Tika, Tonia.

My biggest thanks go to AGDC. You only occasionally made me feel old and you were always up for a brunch at Rickety Press, almost never for burgers at the Gardener's Arms, and very often for drinks at the Jericho Tavern. I honestly couldn't have done it without you. Justin, you are such a joy to be around. Thank you for going to many a concert or comedy show with me. And whenever we inevitably went hiking the same mountains, but separately, I could always rely on a partner and mutual fan of well-documented train journeys and summit selfies. Susie, you are my fastest friend and maybe me only friend equally excited about Park Runs and post-run brunches. Thanks for sitting firmly in the sweet spot of feminism and music, so I always had a keen friend to accompany me. Leila, I'm amazed by your commitment to never miss an opportunity to come and support your friends' endeavours. I am proud of the amount of egg gravy and noodles and gyoza we have jointly diminished. Célia, thank you for going over my ludicrous and probably insanely repetitive sounding French examples and, more importantly, for being the most upbeat company after getting stuck on a train in Germany and missing our flights. Richard, who not only put up with my endless boring requests to please explain HTML or JavaScript to me but who also accompanied me as a mystery diner and would host excellent NYE parties; with Alastair, who will never miss a chance to check in on how his friends are doing. Thank you for always being there to listen and provide an excellent perspective. Adam, I wouldn't know how cricket works without you, and I cannot wait to see you in D.C.!

My friends from other places honestly deserve a badge for all their cheerleading as well as patience while I went AWOL to write these pages. Julia, thank you for always miraculously making time to meet me for lunch or host me when I travel in or out of Hamburg. For coming to see me when I can't get away from Oxford (or Brazil). Tjark, you are one of my oldest friends and reliably awesome. Thanks for being my voice of reason while at the same time entertaining all Star Wars-related plans. Adithya, my always well-informed, open-minded, and ever-supporting friend. I just know that you'll always be there to listen, or help me navigate Beijing airport or the best food in Ann Arbor. Vera, who keeps me grounded and levelled, and will come visit me anywhere to have a lengthy walk followed by a pint, which is the true expression of friendship if I've ever seen one. The Bad Yogi Cool Kids, what a better group of friends to have than those who'll sip wine and eat cheese in France (and do yoga!), and then become the sisterhood of traveling friends. Most importantly, to be a regular entry in the diary during the lockdown and stay-at-home stages of this PhD. Thank you to my family, who have supported me in various different ways and whom I'd be thanking in person by walking Fluse through the woods every morning (if it weren't for the pandemic). I am so grateful and will see you soon! And I couldn't be more grateful to you, Phil (who probably did not know what he signed up for), for not only supporting me in all possible ways during the long days of writing up, but for committing to the combination of getting through lockdown and thesis submission together.

Finally, I thank everyone I ever asked for their linguistic judgements. If nothing else, some people might have stopped believing that what linguists do is speaking many languages (fighting that stereotype one judgement question at a time), and possibly started believing that I do not even speak the one language that we've supposedly been communicating in.

Abstract

How can we account for our extensive knowledge about the world, given that we individually experience very little of it? This thesis considers the mechanisms involved in the acquisition of generic statements. One approach to explain this feat, in light of our limited direct experience of instances of kinds, suggests that language gives us the tools to go beyond our environment. This thesis argues that variation in the morphosyntactic make-up of a generic subject can affect our understanding of how properties are linked to a kind.

One significant challenge for the acquisition of generic subjects is the lack of an overt marker, which means that no generic subject unambiguously encodes genericity. Still, generic subjects can easily be interpreted generically in novel concept acquisition tasks instead of being assigned episodic reference. Experimental research has addressed this puzzle from developmental perspectives Gelman et al. (2010); Prasada (2000); Cimpian et al. (2011) and quantificational ones (e.g. the covert generic operator GEN, Carlson, 1989; Krifka et al., 1995).

The findings are novel as this thesis pays special attention to previously less investigated singular generic subjects. It considers NUMBER and DEFINITENESS and their links to the human conceptual system as the main variables of the make-up of generic subjects. Comparing the morphosyntactic differences of bare plural, indefinite singular, and definite singular subjects experimentally shows that they each raise expectations of tight category-property links, modulated by other available cues.

To better understand the developmental trajectory of generic subjects, this thesis compares adult and child participants' responses to learning about novel kinds. This includes indefinite and definite singular subjects, for which children receive comparatively little input and which are normally deemed highly ambiguous between generic and episodic interpretations.

This thesis concludes that our conceptualisation of novel kinds is dependent on morphosyntactic cues, but crucially, this is further modulated by additional information, such as pre-existing world knowledge, phonological or visual cues.

Contents

List of Figures

List of Tables

List of Abbreviations

AP . . . . Adjective Phrase.

BP . . . . Bare Plural.

ClP . . . . Classifier Phrase.

DN . . . . Definite Numberless.

DPl . . . . Definite Plural.

DS . . . . Definite Singular.

IS . . . . Indefinite Singular.

GaD . . . . Generics-as-Default.

GEN . . . . Covert modal generic quantifier.

GOG . . . . Generic Overgeneralisation Effect.

$\mathbf{L_1}$ . . . . First Language.

N . . . . Noun.

PC . . . . Principled Connection.

quantPL . . . . Quantified Plural.

R . . . . Realisation Operator.

SC . . . . Statistical Connection.

TVJ . . . . Truth-Value Judgement.

#P . . . . Number Phrase.

Much of children's knowledge about the world is obtained not through direct experience but through the testimony of others. (...) Language is a primary means of transmitting information from one generation to the next.

— Gelman et al. (2010) *Effects of generic language on category content and structure*

1
Introduction

Contents

1.1 Motivation

How do we acquire vast amounts of knowledge about the world given that we individually experience so little of it? Even more surprising than the observation that a reader of this thesis has general knowledge about the world is that young children are already relatively apt at understanding what things around them are like. One approach to explain this feat suggests that language gives us the tools to go beyond our environment, to understand sentences that abstract away from daily occurrences and articulate unperceived generalities, and thereby permit us to impart our psychological reality through linguistic knowledge.

To understand whether it is indeed language that enables us to generalise and abstract away, we must first understand the fine-grained linguistic elements alongside the structure of psychological concepts and examine the influence that language exerts over our psychological understanding and vice versa. This thesis will further theories of the human linguistic and conceptual system and address how we humans go beyond our individual experience to understand and express the generalities of the world. To do so, I will address how we acquire world knowledge, the role of language in this process, and more specifically the role of different types of language therein. In other words, I am interested in the **distinct natures of each of the generic subjects**.

The initial observation about generics is that they are widespread and can take various forms. Moreover, these various forms often seem to express the same content, as in the three variations in (1):

(1) a. Triangles are three-sided.
b. A triangle is three-sided.
c. The triangle is three-sided.

All three sentences indicate that all triangles are three-sided. Compare this with the three variations in (2):

(2) a. Mosquitoes carry malaria.
b. ?A mosquito carries malaria.
c. *The mosquito carries malaria.

Only the first version with *mosquitoes* as the subject is felicitous in conveying this (striking) generic. In stark contrast with the assumption that **all** triangles have three sides, the generic in (2) only holds true of less than 1% of mosquitoes. Not only is the corresponding quantified version different, showing that generics can express statements about all members as well as virtually none, but the distribution of acceptable subjects differs as well.

This observation is important in the context that generic statements are ubiquitous and influence human beliefs about the world with ease. Psychology studies conducted at Stanford in the 1970s have become well-known for showing how easily manipulated human reasoning and beliefs are, especially when presented in generic form. For instance, Anderson et al. (1980) investigated the role of discredited information in our judgements in the so-called Risky-Conservative Choice Tests. They demonstrated that, even with very few data points, we generalise and remain confident in recently formed assumptions. While Anderson et al. (1980) were not explicitly researching generic beliefs, their experiments provide supporting evidence for the claim that generic beliefs are easily formed and do not need much, or even any, direct experience with instances in the world. Linguistic research has arrived at similar conclusions. Krifka et al. (1995, p. 72) demonstrated that the phrase *Kim helps her friends in emergencies* "can be true even if there never has been (...) any opportunity for Kim to help friends."

Before I narrow my investigation into our learning about the world and the role of language therein, I would like to outline the relevance of this research from a current societal perspective. It is crucial to contextualise this research further to show the implications of generic statements "in the wild". A rich source of examples that illustrate how the vague meaning and truth-conditions for generic statements lend themselves to be easily misconstrued are online discussions. Consider situations in which an injustice is pointed out and attributed to a social group, or any subset of society that can be stratified. Often, a member of that identified group will point out that this is not applicable to them as an individual (and might retort with "#NotAllMen" or "#AllLivesMatter"). The assumption that a generalisation applies to all members of this (social) group, while generic subjects underspecify the assumed subset (but typically do not presuppose universal quantification), lies at the heart of this argument.[1] A better understanding of the wide range of possible interpretations, based on the acceptable quantifiers generics statements can imply, can aid and mitigate these misunderstandings.

In more precise terms, this can be specified as targeting the role of language in acquisition of (abstract) world knowledge. This has long been a multidisciplinary effort, with a multitude of approaches in the study of generic statements. One contrast within the domain of generalising statements emphasises the difference between generalising statements and kind-referring ones, relying on the notion that

[1]Similarly, for a less academic take on why these are problematic when taken as universally-quantified statements rather than generic generalisations, see Strickland (2017).

"with kinds we abstract away from particular objects, whereas with characterizing sentences we abstract away from particular events and facts" (Krifka et al., 1995, p. 4). This conceptual distinction has significant bearing on the overall field and cannot be dissociated from the linguistic differences between generic as opposed to specifically kind-referring subjects.

Seeing that linguists, psychologists, and philosophers have long been considering how world knowledge is acquired, especially when our direct experience of (instances of) kinds is highly limited, what is the role of this thesis? My primary contribution is a more unified response to the question of how we acquire and use generic reference. We will see that in the absence of an explicit marker of genericity no generic subject can unambiguously encode it. Still, generic subjects are easily (and typically) interpreted generically in novel concept acquisition tasks instead of being assigned episodic reference. For this more interdisciplinary response to be applicable to a wide range of data, I will take seriously the differences between generic subjects that have been defended in the linguistics literature. Thus, the theoretical part of this thesis focuses on the role of morphosyntax of the subject DP, considering the linguistic variables NUMBER and DEFINITENESS per se as well as their links to the human conceptual systems. This relies on the assumption that the various morphosyntactic make-ups of each generic subject affect our understanding of how certain properties are related to the kind in distinct ways.

1.2 Knowledge about the world

Even though the questions that linguists, psychologists, and philosophers have asked target slightly different specifics, they all aim to understand how we make sense of the world. The underlying motivation is that our knowledge is abstracted from individual instances or stories about individual instances in temporal-spatially bound events, and (in an opaque manner) transformed into knowledge about kinds and categories. I will address ideas about the nature of kinds below but this is the overarching question that will guide the more specific linguistic and psychological questions about the nature of generic statements. These more specific questions are an attempt to break down the significant question of how we come into possession of world knowledge.

The question of how we come to know anything about the world is not new. The humanities and social sciences in particular have dedicated much theoretical and experimental effort to investigate how the brain processes conceptual knowledge. Philosophers have addressed the “problem of induction”, which is intricately linked to wider philosophical questions (e.g. the nature of human a priori reasoning and cause-and-effect therein, as in Hume (1748, esp. p. 330-332)). This thesis therefore joins a long line of research investigating how we confidently make vast generalisations when appropriate, yet refrain from it when they are not licensed (even when they could be acceptable from a logical perspective). Following this tradition, many theories have been proposed (and consequently discarded). One such theory is the Sapir-Whorf hypothesis which gained traction in the early 20^{th} century and falls under linguistic relativism, advocating for linguistic determinism. This theory of linguistic determinism argues for an extreme version of how language determines our perception of the world (e.g. Sapir, 1921; Weisgerber, 1950; Whorf, 1956). A *weak* version of this hypothesis argues that language influences our perception of the world, is still popular.[2] Rather than assuming that language is the sole determiner in our perception of the world (à la a “language as the window to the world” account), I will explore via a series of experiments to what extent it influences perception of category knowledge, and, crucially, which other factors must be considered.

Our world knowledge covers many aspects of the world, albeit to different extents. For example, a microbiologist’s knowledge about microbes will be much more detailed than mine. This thesis is explicitly concerned with everyday expectations of the objects and individuals surrounding us rather than specialist knowledge. The first set of studies leaves the type of objects, and therefore the kind itself, underspecified. The subsequent experiments will focus on animal kinds for a variety of reasons

[2] The *strong* hypothesis of linguistic determinism recently came to mainstream fame by featuring in the 2016 sci-fi movie *Arrival*, arguing that the language we speak determines not only our perception of the world but also our linear understanding of time (based on Ted Chiang’s 1988 short story *Story of Your Life*).

outlined in each of the experimental chapters. At this point, I would simply like to note that they differ from other kinds, most notably by having higher essentialism than artefacts (see e.g. Gelman et al., 1998). This is further nuanced in Prasada & Hall (2019)'s discussion of transformation theories between different basic level and subordinate kinds. They mention that "members of natural kinds [are] less mutable than members of artifact kinds." (2019, p. 212) With this statement, they also allege the high essentialism that is associated with animal kinds.

Ultimately, language plays a crucial role in passing on world knowledge from one generation to the next. Yet, the role of linguistic cues must be seen within the context of other available cues, as even this brief example of differences in the level of essentialism between animal and non-animal categories types illustrates. I will analyse the role of linguistic cues in detail and in different contexts. I aim to establish its influence on acquisition of world knowledge, to ascertain that genericity should not be treated as one monolithic linguistic phenomenon but under an analysis that reflects its varied semantic contributions to the conceptual system.

1.3 Generics and truth-value judgements

One final observation that motivates my research is the unique status of generic statements when it comes to their truth-value judgements and correspondence to subsets. One way in which the difference between generic and episodic statements becomes apparent is that generics resist contextual restriction. Focusing on comparing the conditions under which these two types of phrases are felicitous will hopefully lead to a better understanding of the nature of genericity. Take, for example, the following statements about tigers in the zoo, raised by Lazaridou-Chatzigoga & Stockall (2013, p. 326, their example (9)). The crucial difference is that even if you see specific tigers in the zoo, by making a generic statements with the subject *tigers*, it will be interpreted about tigers in general and not the contextually salient tigers:

- Context: There are lions and tigers in this cage.

[a] Every lion is dangerous.

(*Can* mean ‘Every lion in this cage is dangerous’)

[b] Lions are dangerous.

(*Cannot* mean ‘Lions in this cage are dangerous’)

In more general terms, much of the initial research into the nature of genericity has looked at the more readily observable evidence that generic language is unlike quantified statements, as illustrated here by the difference between *every lion* and *lions*.[3]

Comparing statements with an overt quantifier such as *some*, *all* or *few* unveils the difficulties in acquiring the meaning of generic statements, which can be hypothesised to be due to the lack of a clear correspondence between one linguistic form and one subset relation. This lack of truth-value conditions has puzzled many semanticists. However, developmental psychology has provided a theory under which generic statements are acquired early and effortlessly, namely arguing for *Generics-as-Default* (see among others Gelman & Raman, 2003; Gelman, 2004; Hollander, 2007; Leslie, 2008). Under this account, generic reference is the conceptual default and thus need not be acquired. Instead, it is specific reference and quantifiers that challenge our conceptual system. This has been explicitly argued for bare plural subjects, in e.g. Leslie (2008).

In this introduction, I have outlined general observations and considerations about generic statements and how they are used in everyday language. They have continued to pique interest in linguistics and psychology as well as philosophy and cognitive sciences because of their versatile and ubiquitous nature. Gelman (2004, p. 6) concisely summarises the main problem, i.e. that “[s]imply put, there is no one-to-one mapping between form and meaning, in the case of English generics.”

The next chapter will review theoretical questions by looking at the contributions of DEFINITENESS and NUMBER in generic subject DPs. This builds on earlier work

[3]If the difference between singular and plural number seems to be an issue, the example works with *All (the) lions are* as well, although judgements seem to be subject to more variation in examples without the definite determiner.

on generics (e.g. Lawler, 1973; Krifka et al., 1995), but also more recent cross-linguistic investigations (e.g. Borik & Espinal, 2012, 2015; Borik, 2016; Dayal, 2004). The subsequent chapters present a series of experiments that make use of various paradigms that have already been applied in studies on concept acquisition to test whether the results corroborate the theoretical assumptions and linguistic distinctions set out in Chapter 2. These studies address the role of varying amounts of exposure to linguistic cues, how different types of cues interact with linguistic cues, whether different types of categories rely more heavily on linguistic cues or whether they are more resistant to morphosyntactic variation, and finally developmental trajectories within the context of genericity and acquisition of world knowledge.

*Guy who invented **the photograph**: I invented **the photograph**!*
*Guy who hates the guy who invented **photographs** and is about to invent **puzzles**: Cool can I see that?*

— @TheAndrewNadeau *Twitter, Jan 4, 2018*

2 Concepts, Properties, and Language

Contents

The goals of this literature review are threefold. First, this chapter aims to uncover the links between the syntax-semantics interface in generic subject DPs, emphasising the differences between bare plural, indefinite singular, and definite singular subjects. Second, this review will highlight the role of and interaction between the various generic subjects and their properties. Third, it will provide a more detailed insight into the type and amount of input of generic subjects that children receive during language acquisition. The sections of this chapter will follow this structure and order. More specifically, Sections 2.1-2.4 build the nominal spine of the generic subject DP. I will start by exploring its smallest component, i.e. the noun itself, and build up the covert and overt layers of the DP. This should aid our understanding of the distinct structures of the three generic subjects of count nouns in English. Spelling out these differences in detail is important for the hypothesised interpretational and behavioural differences in linguistic judgements, which are introduced in Sections 2.5 and 2.6. This literature review will end on looking at child-directed speech in two ways: First, as collected in two of the CHILDES database corpora and second, as presented in two picture books. This enables us to better understand the input that children receive at various stages of the first language development.

Before I present the linguistic data that serve to exemplify the noun denotations and the various properties of generic subjects, note that these are based on introspective linguistic judgements. Future work would benefit from collecting empirical data, specifically designed to gather grammaticality judgements about the different types of generic statements. Seeing that thesis seeks to address the gap between approaches in linguistics and in psychology, collecting this type of empirical data as the basis for evaluating the strength of the assumptions would address one of these gaps and help strengthen hypothesis formation. However, note too that linguists have assessed the reliability of linguistic judgements, notably Sprouse & Almeida (2012) (specifically focusing on the criticism that informal judgement collection leads to false positives) and Sprouse & Almeida (2017) (with a focus on possible false negatives). This and subsequent work has found that “traditional methods are a highly reliable method in their own right” (Sprouse & Almeida, 2012, p. 633). Sprouse & Almeida (2012) found at least 98% replicability of textbook judgements in a magnitude estimation task, i.e. a task in which participants rate the acceptability of a sentence with reference to a 'standard' sentence that is assigned an acceptability rating (called the 'modulus'). In Sprouse & Almeida (2017), they followed up by using a combination of four types of grammaticality judgement tasks, adding yes-no, two-alternative forced-choice, and Likert scale tasks to the previous task of magnitude estimation. While the tasks differ in terms of sensitivity, the overall conclusion remains that "acceptability judgment experiments with relatively small sample sizes [i.e. informally collected judgement data] are actually relatively well powered to detect many of the phenomena currently in the syntax literature" (Sprouse & Almeida, 2017, p. 28).

2.1 Noun denotations

It is widely accepted that the interpretation of a statement as generic depends on combining a variety of cues (e.g. Krifka et al., 1995; Kratzer, 1995; Asher & Morreau, 1995). In this thesis, I am primarily concerned with generic subjects and therefore NPs and DPs. At the heart of this lies the interpretation of a noun. Thus, it only seems fitting to begin by giving an account of the denotations of nouns that I assume for the subsequent analyses. I will provide an overview of the two main accounts that have been defended in linguistic theory, namely 'nouns as kinds' and 'nouns as properties of kinds', which I will explore in Section 2.1.1 and 2.1.2, respectively.

I begin with these to provide a baseline of my understanding of the meaning of nouns as they are at the core of conveying the meaning of generic phrases. Of course, 'understanding the meaning of nouns' is an undertaking of a scope that deserves its own thesis. The purpose of this section is to provide an overview of the various approaches at the interface of syntax and semantics, which provides a working definition that allows us to understand the motivations for the experimental chapters that follow.

This brings us to the second and closely related aim of this section, namely to understand the relationship between NPs and genericity. Consider Topolinjska (2006, p. 55), who argues that there are no intrinsically generic NPs, or NPgen as she calls them, but only NPs with generic usage. NPgens, according to Topolinjska, refer to a genus, which is defined as a concept with non-accidental, inherent features. In the discussions to follow, I will reinterpret this specific type of NP as a directly kind-referring NP, rather than a generic NP. She argues that if an NP makes reference to a genus, then this NP has intensional content which is applicable to the kind (although more on that distinction in Section 2.6). The distinction between kind-referring and generic subjects is closely linked to a proposition put forward by Prasada & Dillingham (2006, et seq.), differentiating between principally connected and statistically connected properties of a kind.

Alternatively, a generic NP might rely on existential presuppositions, and thereby does not have intensional but extensional content. One undesirable consequence of this understanding of generic NPs is that hypothetical and counterfactual kinds could not be accounted for in a straightforward manner, although this could be mitigated by stipulating that "there is no presupposition about the existence of any specified denotates" (Topolinjska, 2006, p. 58). This distinction between kind reference and generic reference seems particularly defensible considering that Topolinjska argues that clear-cut distinctions regarding category membership can be easily made by stipulating NPgens vs NPs with generic usage.

This distinction can be boiled down to the distinction between the two theories, intension vs denotation, which seems to target the difference between kind and generic reference. Ultimately, I argue for a distinction between non-generic, generic, and kind-referring NPs based on differences in the syntactic spine of the DP. In other words, I will rely less on semantic differences in the nouns denoting the sentence's subject than on the differences in their syntactic layers, although these inevitably impact the tight syntax-semantics interface.

Before looking at the syntactic layers, let us focus more on the two main conceptions of noun denotation: Nominals can be seen as naming kinds directly, an idea I will explore in Section 2.1.1 or as denoting properties of kinds, as explained in Section 2.1.2.

2.1.1 Nouns as kinds

The first conception of the denotation of nouns is that they are names for kinds. Under the most famous analysis of this kind, the Carlsonian theory, nouns are descriptors and receive a uniform analysis. To account for episodic reference, they can be shifted, or converted, into referential expressions via the realisation operator **R** (Carlson, 1977). The assumption that nouns are names for kinds feeds into an ongoing debate, with particular relevance to modified kinds (which I will elaborate on in Section 2.3.3). To address the debate regarding domains, i.e. names for kinds as opposed to names for objects, Carlson assumes that all bare plural (BP) subjects in English should receive uniform treatment and that they are "all occurrences of the same syntactic or semantic structure" (1977, p. 3). Consequently, all differences in interpretation are due to differences in the context.[1] Carlson (1977, p. 61) proposes that English BPs, as descriptors, are names for kinds of things, not objects themselves. This approach seems favourable, but subsequently likens generic BP subjects to names for individuals (1977, p. 78). This means that the difference between a generic statement and existential one is as follows: *Dogs run* translates to the set of properties denoted by *dogs*, while *Dogs ran* differs in two ways: both the existential quantifier ∃ and the realisation operator **R** are applied for the existential statement. This interpretation can maintain a unified analysis of the BP subject *dogs*.

This discussion and many subsequent ones based on Carlson are based on either the semantics of nouns in general, or on the interpretations of generic NPs. By taking nouns to denote kinds directly, differences in the covert syntactic layers can explain interpretational differences both between various types of generic subjects and generic and non-generic subjects. This approach is supported by investigating a much wider range of generic subjects than only English BP subjects, as well as a

[1] Carlson (1977, p. 22) further assumes that there is only one type of generic usage, while this thesis distinguishes between multiple types of generics, following e.g. Leslie (2008); Lazaridou-Chatzigoga et al. (2019), inter alia.

wider range of types of generics.[2] Only by considering a wider range of linguistic data, both in terms of types of generic statements and cross-linguistic patterns, can the hypothesis that nouns are names for kinds be tested. Even so, it seems possible to stipulate a kind-referring root to nouns and attribute differences in interpretation to other factors (see also Borer, 2005).

Another account that takes nouns to be names for kinds is given by Schmitt & Munn (1999). They defend the importance of NUMBER in arriving at various interpretations of nouns. Schmitt & Munn draw upon data from Brazilian Portuguese, which has both a SINGULAR/PLURAL distinction and a MASS/COUNT distinction. Normally, this implies that individuals are referred to with their appropriate NUMBER value, if the noun is a COUNT noun. Yet, there are cases in which a NUMBER-neutral noun is used in generic sentences, and anaphora is only allowed via a null-pronoun or a plural one. Schmitt & Munn (1999, p. 354) summarise that "the locus of crosslinguistic variation lies in the interaction between the determiner system and the morpho-syntax of Number", with further reference to Dayal (1992). This observation motivates the consideration of the role of determiners rather than the pursuit of an analysis of the meaning of nouns without additional layers of the DP.

Regardless of the specifics of Carlson's or Schmitt & Munn's proposals, an additional upshot of this theory is that other phenomena that we observe in language acquisition cross-linguistically, such as labelling effects, could easily be accounted for. When introducing novel objects in experimental settings, some researchers (e.g. Gelman et al., 2010) have compared not only generic labels with specific labels (such as *zarpies* vs *this zarpie*), but also compared using a label at all as opposed to no label and whether that aids categorisation and memory (i.e. comparing either one of the former with only using *this*). Developmental psychologists have consistently found that using a label for an object enhances responses with regard to category-property links and ratings of essentialism in children and adults. Therefore, if

[2] This was more recently defended by Nguyen (2019), although this account excludes *k-level* predicates and therefore focuses on generic but not kind reference. Further, Nguyen concludes that, ultimately, BPs are semantically incomplete and do not express a proposition, which is almost diametrically opposed to the perspective defended here.

we assume that nouns themselves denote kinds, then by using a noun as a label, reference to a kind is automatically made. Consequently, establishing that a kind exists raises expectations that a category structure exists, which has been argued for independently by Topolinjska at the start of this section.

2.1.2 Nouns as properties of kinds

The second main conception is that nouns are properties of kinds, and denote kinds once they combine with a definite determiner. McNally (2017) aptly illustrates that the inverse of the well-known Carlsonian approach to the meaning of nouns, i.e. not names for kinds but nouns as descriptors, as introduced by Zamparelli (2000). He argues that nouns are referring expressions first, and can be converted into descriptions. Unfortunately, even if we take nouns to be converted to descriptions, this still allows for them to be descriptions of kinds (as well as of entities) and does not settle the debate at hand.

Let us look at an elaborated version for the kind-referring properties of nouns. Borik & Espinal provide a minimal pair of the denotation of nouns compared to definite NPs. They explicitly argue for the denotation of nouns as properties of kinds.[3] Crucial for this distinction in the analysis of kind-referring subjects (which Borik & Espinal refer to as definite numberless (DN), a theory of which I will explore in more detail in Section 2.3) is the generalisation that if a language has determiners, it has definite kinds. The N as the nominal root is kind-denoting (1) and the function of ι is to bind variables of kinds (2) (based on Borik & Espinal, 2012, p. 131):

(1) a. $[\![\mathrm{N}]\!] = \lambda\ \mathrm{x^k}\ [\mathrm{P(x^k)}]$

b. $<\mathrm{e^k}, \mathrm{t}>$ [properties of kinds]

(2) a. $[\![\mathrm{D\ N}]\!] = \iota \mathrm{x^k}\ [\mathrm{P(x^k)}]$

[3] The theory of common nouns as functions denoting properties has also been developed in contexts such as possible world semantics. See e.g. Oltean, 2009, p. 269, referring to Fregean *Sinn* and *Bedeutung* and Kripkean rigid designators, arguing that common nouns have the following function, exemplified by the noun *dog*: $[\![\mathrm{dog}]\!] = \lambda$x. *x* is an entity. *x* is a dog. I will not discuss these interpretations at length here as rigid descriptors are normally taken to denote a set of individuals and this thesis is particularly concerned with the interpretation of kinds, which might not have any members.

b. <e^k> [kind denotation]

In this context, this approach is a preferable framework. This is mainly due to the fact that it can explain how DEFINITENESS enables reference to kinds, in languages with and without overt determiners as opposed to INDEFINITE, generic reference. Separating how a subject makes generic vs kind reference is split by the presence of D under this theory. Under theories that argue that nouns are names for kinds, such as Carlson's, both the role of D as well as the shift from kind reference to generic reference, while maintaining the evocation of non-object level reference would be more difficult to account for. The exact role of the determiner will be explained in Section 2.2 below. However, these types of theories also tie in well with upcoming notions of NUMBER, to be discussed in Section 2.3. Krifka (2003, p. 185), referring to Chierchia (1998), also treats nouns as "a function that maps every world w to the set of (atomic) dogs in w", and gives them the following denotation:

(3) $[\![\text{dog}]\!] = \textsc{dog} = \lambda w \lambda x[\textsc{dog}(w)(x)]$

Ultimately, these types of accounts offer more flexible interpretations cross-linguistically and in terms of readings with wide and narrow scope, while maintaining, at its core, reference to kinds due to the nature of the nominal.

To conclude, this section has introduced two different theories about noun denotations. For the purpose of this thesis, it is less crucial whether the noun itself returns a kind directly or properties of the kind. As an aside, this is in line with theories that, cross-linguistically, nouns may have different denotations. One such example, based on Chierchia (1998)'s Nominal Mapping Parameter, subdivides languages into [±arg] and [±pred]. Based on their parametric settings, noun denotations differ between languages. More importantly, both theories stipulate that nouns are, in one way or another, closely related to naming kinds. In order to contribute to the conversation on generic subjects in a meaningful way and explain phenomena observed in psychology, such as *Generics-as-Default*, as well as in semantics, such as stipulating GEN, my main concern is to assume that by using a nominal in a generic subject, reference to a kind and/or its properties is made.

2.2 Determiners

The previous section has argued that nouns can be analysed as names for kinds or as properties of kinds. Following the latter approach, the question now is to understand the role of the other components of generic subjects. I assume that apart from the content conveyed by the N itself, semantic information can be encoded in the structure of the DP (specifically following Borer, 2005, p. 68, but see also Longobardi, 2001; Alexiadou et al., 2007; Alexiadou, 2014). I will work my way up the layers of the subject DP. The purpose of this section is to explore the function of the definite determiner in kind-referring subjects in English, the indefinite determiner, and an absence of determiners altogether, accounting for minimal pairs as in (4):

(4) a. The tiger has stripes.
  b. A tiger has stripes.
  c. *Tiger has stripes.

This example shows that bare count nouns are ungrammatical for kind reference in English. A discussion on exceptions to the obligatory use of definite determiners within the context of German generics follows in Section 2.2.1.

I assume kinds not to be individuated. Individuation, or divisibility, is introduced by the Classifier Phrase ClP. Note though that the two determiners, *the* and *a*, do not fulfill the same function. This implies that the underlying form of (4-a) differs from that of (4-b). While they seem similar, they do not allow for the same patterns of co-occurrence with number markers (5), plural markers (6), and countability (7):

(5) a. The one tiger.
  b. *A one tiger.

(6) a. The tigers.
  b. *A tigers.

(7) a. The flour.
  b. *A flour.

These distinct distributions point to different underlying structures, which distinctly have to do with individuated objects, especially based on examples such as (7). I therefore follow Borer (2005, p. 110), and propose that both the indefinite determiner and plural marking sit in ClP (8), based on their complementary distribution, while the definite determiner sits in #P (9):

(8) $[_{\#P}$ $[_{ClP}$ a $[_{NP}$ tiger]]]

(9) $[_{\#P}$ two $[_{ClP}$ tigers $[_{NP}$ ~~tiger~~]]]

The notion of DEFINITENESS, especially when discussing English definite subjects, needs to be explored within the context of previous semantic approaches of interpreting definite determiners. Semantic work previous to Schwarz (2009) has mostly focused on one of two presuppositions that are necessary for DEFINITE reference: *familiarity* and *uniqueness*. The notion of uniqueness was especially prevalent in determining semantic operators. In this spirit, semanticists such as Bierwisch (1970) assume that there are different types of semantic operators applying to lexical items. More recently, Coppock & Beaver (2015) discuss the uniqueness presupposition for English (weak) determiners, independent of an existence presupposition, which is an important distinction to bear in mind for the discussion of definite singular (DS) kind-referring subjects (see p. 29).

Even though the feature DEFINITE is usually expressed by the definite article in languages with articles, e.g. *the* in English, this is not necessary for languages without determiners (as we will see in the discussion of Slavic examples, and Polish in particular, see p. 43f). DEFINITENESS is often related to definite descriptions of the type the F, which are semantically expressed by the iota-operator (ι). Definite descriptions presuppose a unique individual or object that corresponds to the description as in the famous example *The present King of France is bald* (Russell, 1919). While this leads to interesting problems, such as how to interpret a sentence about an individual that does not exist, as France has no king (Strawson, 1950; Frege, 1948), we are more concerned with the interpretation of DEFINITENESS as expressed by ι. Partee (1987) first brought attention to the definition of ι, followed by many others, e.g. Chierchia (1998). Partee argues that ι is (at least in its referential use) entity-denoting and is a partial, surjective function. If applied to the denotation of N, it renders the unique x. In English, this is implicit in minimal pairs such as those in (10-a)-(10-b) (Partee, 1987, p. 125, her (12). NB: This is generic reference in predicate position but the minimal pair in (4) showed similar differences for generic reference in subject position.):

(10) a. John is the president / president.
b. John is the teacher / *teacher.

Interestingly, as is it presupposed that there is one and only one president (within the relevant context), the denotation of the definite determiner need not be made explicit under these circumstances (10-a). However, no presupposition about 'at most one teacher' is salient so that the is necessary in the predicative use of teacher in (10-b). Chierchia's contribution in explaining ι elaborates on Partee's definition by applying it to sets of individuals, both as 'singularities' if combined with a singular noun as in (10), but also as 'pluralities' if combined with a plural noun (Chierchia, 1998, p. 346, his (11)):

(11) a. ι X = the largest member of X if there is one (else, undefined)
b. the dogs = ι DOGS = the largest plurality of dogs
c. the dog = ι DOG = the only dog (if there is one)

Both (10) and (11) show that ι presupposes at least one atomic instance of an individual.[4] Additionally, (11-c) allows for group-readings which contrast with generic statements made with BP subjects.

Having explored some definitions of DEFINITENESS, let us move on to the link between DEFINITENESS and the morphosyntactic layers of the generic DP. The definite determiner captures, at least for English-type languages, how reference to kinds is established. Crucially, this is not possible with bare nouns (e.g. for *tiger* in

[4] Note that the application of the definite determiner under this interpretation is compatible with mass nouns. I assume mass nouns to be similar to plural and therefore ι can readily apply to them as it does to plural count nouns.

(4-c)). This contrasts with languages without overt definite determiners, which can still have definite kinds, where DEFINITENESS is encoded by a null-determiner. By focusing on the acceptability judgements of generic sentences with definite subjects vs those with other kinds of subjects, note the relevance of the hypothesis that definite subjects denote the kind itself (following Borik & Espinal, 2012). In Section 2.3 on the role of NUMBER, data from both Spanish, where reference to kinds requires the definite determiner, and Russian, which lacks determiners altogether, will exemplify the significance of both DEFINITENESS and NUMBER for kind reference.

In addition to Borik & Espinal, arguing that nominals denote properties of kinds, and that definiteness makes the NP kind-referring, one could also adopt a morphosyntactically complex view of determiners as introduced by Leu (2015). His treatment of kind-referring NPs leads to the problem of how we analyse its non-kind-referring counterpart, i.e. the *definite singular* used in specific or episodic contexts. One way to address this ambiguity is to stipulate a homophonous determiner, such as *the* in English. According to Leu's theory, determiners are complex and can be decomposed. Syntactically, they form part of an extended AP, namely xAP. Interestingly, by stipulating that certain determiners are in fact demonstratives, and therefore part of a complex phrasal structure, they have the syntactic set-up to include a projection for silent deictic elements. Leu specifies them as, e.g. *here* or *there*. With underlying syntactic differences between articles and demonstratives, phonological alternations can be accounted for not just as idiosyncratic but rule-governed alternation. The relevance of this alternative account will become more obvious when comparing optional determiners in generic phrases. To do so, I will now provide an overview of generic subjects and how they interact with mass and count nouns.

2.2.1 Mass and count nouns

If we assume that the life of the nominal starts out as referring to properties of kinds (but recall views on nouns as names for kinds have been discussed in Section 2.1.1), then we need to be able to account for reference to instances of said kinds. Let us begin by looking at individuals which can be divided into two types: nouns for mass individuals and nouns for count individuals. Let us bear in mind this difference to better understand the relevant layers of the DP and their function in the domain of kinds as well as of objects. Individuation as a concept is tightly linked to the structures of the DP, which may be covert. Particularly, we pay attention to the detailed analyses laid out by Borer (2005, 2009), and Borer & Swart (2011). Following Borer, individuation for nouns is introduced by ClP, which introduces countable structures. Thus, a noun that has a mass interpretation must lack a ClP. Under this account, all nouns start out with a mass interpretation and are individuated and countable through this additional syntactic layer (although see Cheng et al., 2008 for a criticism of this theory, related to the data motivating Borer, 2005's syntactic theory and claiming that if a 'universal grinder' exists, it at least cannot be universal.)

An alternative way of accounting for differences between mass and count interpretations is Lewis' 'universal grinder' theory (Lewis, 1975, as cited in Pelletier, 1975). While the specifics of Borer's and Lewis' arguments are distinct, the main idea is the same. If a (possibly) countable noun lacks a plural marker (which is part of ClP according to Borer's theory), the noun is interpreted as mass as it has undergone the process of the universal grinder. Pelletier (1975, p. 456) elaborates: "The reader has doubtless guessed by now the purpose of our universal grinder: Take an object corresponding to any (apparent) count noun he wishes (e.g. 'man'), put the object in one end of the grinder and ask what is on the floor (answer: 'There is man all over the floor')."

On the other hand, Pelletier (1975, p. 456) points out the intuition about subkind references with plural mass terms, exemplifying that " 'How many oatmeals are in your kitchen?' (...) can be given a perfectly clear count sense. This sense might be the same as that of 'kind of oatmeal'..." This connection between countability, subkind reference and NUMBER is also defended here. To defend the existence of a ClP, we must then acknowledge these systematic differences between using a determiner with a mass term or using one with a count noun.

These somewhat divergent uses show an alternation of countability (as well as accessibility of subkind interpretations).

One particularly interesting case concerns alternations in the presence of determiners for mass terms. In German, generic subjects for mass terms typically occur without a determiner (13-a), as they do in English (12). However, the definite determiner may optionally be present in some generic expressions (13-b):[5]

(12) Gold is a precious metal.

(13) a. Gold steigt im Preis.
Gold increases in.the price.

b. Das Gold steigt im Preis.
The gold increases in.the price.
'The price of gold increases.'

Within this context, recall Leu's theory which might be able to account for this alternation. Specifically, the phonologically stressed determiner, *-d*, realised in German as {*der, die, das*}, may not occur in kind-referring subject DPs (14-a). Only the unstressed variant (14-b), which Leu analyses as an article residing in the D-head and not a demonstrative (and it thus does not have a silent deictic element) may occur in these instances:

(14) a. ***Das** Gold steigt im Preis.
The gold increases in.the price.

b. Das **Gold** steigt im Preis.
The gold increases in.the price.
'The price of gold increases.'

[5]Note that the presence of the definite determiner is more restricted for kind-referring subjects that co-occur with kind-level predicates. While (i-a) is extremely common, (i-b) is more archaic and less used nowadays (judgements are my own, (i-b) is taken from Sallmayer (1865), and cross-referenced with Ngrams on Google Books):

(i) a. Gold ist ein Edelmetall.
Gold is a precious.metal

b. ?Das Gold ist ein Edelmetall.
The gold is a precious.metal

Instead, this emphasis implies the kind of specific reference a strong demonstrative carries, as shorthand for *This (piece of) gold* having a certain property.

Gunkel et al. (2017, p. 332) also argue that for count nouns, the definite determiner is often optional in German. This has been analysed as part of the grammaticalisation of the definite determiner in German (adapted for legibility from Hawkins, 2004, p. 85, further supported by Krifka, p.c. at SPE10): "German has gone further than English and regularly uses the definite article with generic plurals where English does not [emphasis and gloss added]:

(15)

Er	zieht	den	**Rosen**	die	**Nelken**	vor
he	prefers	the.DAT.PL	roses	the.ACC.PL	carnations	PREFIX

'he prefers carnations to roses'

He prefers the carnations to the roses in English suggests pragmatically identifiable sets of each." Other examples confirm that k-level predicates easily co-occur with plural generic subjects:

(16)

Die	Wölfe	sind	am	Aussterben	in	diesem	Land
The	wolves	are	on.the	verge.of.extinction	in	this	country

'Wolves are on the verge of extinction in this country.'

Gunkel et al. continue that out of a set of eight major European languages that have covert definite determiners (German, French, Italian, Romanian, Hungarian, Greek, English, and Dutch), German is the only language that can optionally employ the definite determiner in generic contexts. Neither English nor Dutch allow their respective definite determiners, while the others require it. Note how this paradigm is reversed once abstract nouns are referred to generically (Gunkel et al., 2017, p. 335, their (148)):

(17) a. (*The) Life is complicated.

b. *(Das) Leben ist kompliziert.

The	life	is	complicated.

'Life is complicated.'

These examples show that any theory of DEFINITENESS and definite determiners needs to account for alternations, which ultimately seem to interlink with phonological cues. These can also help represent the different semantic components of DEFINITENESS (which include familiarity and uniqueness). However, why this distinction between abstract and concrete nouns exists continues to be unclear to me.

This section has provided a detailed overview of the contribution of definite determiners and the role of the ι operator for subject DPs. This is relevant to the research pursued in this thesis as it investigates the morphosyntactic pieces in generic subjects and how their different semantics raise distinct expectations in the human conceptual system. Without these semantic differences, expressed in the syntax, generic subjects should be expected to behave the same. Yet, their distribution patterns show that this is not the case. This will be considered in the analysis of the experimental data in later chapters.

For now, the other variable that is relevant to distinguishing the three generic subjects for count nouns in English is NUMBER. I will present the morphosyntactic differences between singular and plural NUMBER in the following section.

2.3 NUMBER

After looking at definiteness and its role for uniqueness, this section motivates additional layers in the DP, which account for NUMBER, and more specifically the position that the plural marker takes in English. Recall the assumption that semantic content can be expressed by syntactic layers of the DP. In Section 2.2, I illustrated the necessity of the definite determiner in English with the minimal pair *The tiger has stripes* and **Tiger has stripes*. In a parallel fashion, the role of NUMBER for count nouns can be illustrated by the following minimal pair:

(18) a. Tigers have stripes.
b. *Tiger have stripes.

I expect that languages will mediate the semantic properties related to NUMBER through the functional projection #P, based on Borer (2005).

One major initial concern that is often raised regarding NUMBER, and PLURAL especially, is the role of such a DP layer in languages without plural morphology.

In these cases, we can draw a close analogy with classifiers. They are similar to the plural in that both are required in the presence of count nominals. Classifiers are necessary when nouns are quantified and when they have a countable structure as opposed to a mass denotation. In other cases, classifiers are not present. Let us begin by focusing on these cases to illustrate that these close parallels exist. Since classifiers function as independent elements, classifier-languages such as Chinese need a projection for the classifier-head (adapted from Borer, 2005, p. 86):

(19) liang ge ren
two CL person
'two people'

Note that there is no plural marking on the N 'ren' like there would be in English. The classifier and plural are similar in that both are required in the presence of count nominals. Compare to the English version *One person* vs *Two people*, or, to use a regular plural, *One tiger* vs *Two tigers*.

We find an important piece of the puzzle that links the similarities between plural and classifiers in languages such as Armenian, which has both a plural morpheme and a classifier system. Interestingly, plurality for count objects can either be expressed via plural morphology on the N (20-a), as in English, or via a classifier (20-b), as in Chinese. Crucially, the two cannot co-occur (20-c) (as in Borer, 2005, p. 94):

(20) a. yergu hovanoc-ner uni-m
two umbrella-PL have-1SG
'I have two umbrellas'

b. yergu had hovanoc uni-m
two CL umbrella have-1SG
'I have two umbrellas'

c. *yergu had hovanoc-ner uni-m
two CL umbrella-PL have-1SG
'I have two umbrellas.'

It is a standard methodological assumption in linguistics that if two items cannot co-occur, they likely compete for the same functional slot. As Borer lays out concisely, "[i]f we adhere to the age-old practice of assuming that complementary distribution is an indicator of structural competition, it follows that [the two items] are, at least in some sense, competing for the same structural slot" (2005, p. 35). In optional classifier-languages such as Armenian, only one mechanism to mark plural can be present at a time.

The examples reviewed above establish the need for a functional head between D and N, to host elements related to NUMBER. This may be a classifier, emerging when we count things. Hence, we need to extend the DP to include another intermediate functional projection. In syntactic terms, we will assume the following syntactic projections (as also introduced by Heycock & Zamparelli, 2005; Borer, 2005; Acquaviva, 2008, among others):

(21) $[_{\text{DP}}$ D $[_{\#\text{P}}$ Num $[_{\text{ClP}}$ Cl $[_{\text{NP}}$ N$]]]]$

where D is projected in the Determiner Phrase DP, providing definiteness and hosts definite determiners, quantifiers, and demonstratives (following Alexiadou, 2014, p. 9); #P is the NUMBER-phrase, providing plurality and hosts cardinal numbers and special numbers (following Acquaviva, 2006, 2008); ClP is the Classifier Phrase, providing divisibility and the mass/count distinction and hosts indefinite determiners and PL-marking (following Borer, 2005, p. 93); and NP hosting the lexically contentful noun. To illustrate some of these differences, see the following example by Borer (2005, p. 110):

(22) a. boys = $[_{\text{DP}}$ $[_{\#\text{P}}$ $[_{\text{ClP}}$ boys $[_{\text{NP}}$ ~~boy~~$]]]]$
b. ten boys = $[_{\text{DP}}$ $[_{\#\text{P}}$ ten $[_{\text{ClP}}$ boys $[_{\text{NP}}$ ~~boy~~$]]]]$
c. salt = $[_{\text{DP}}$ $[_{\text{NP}}$ salt$]]$

With the functional projections of NUMBER introduced for reasons independent of genericity, the question is how such a functional projection relates to the distinction of generic vs individual readings, i.e. (23) vs (24).

(23) The tiger has stripes. [generic/individual]

(24) This tiger has stripes. [*generic/individual]

Based on the assumed denotations for nouns (Section 2.1.2, especially p. 15), I propose that the kind reading in (23) does not have a #P projection, and is therefore not numbered singular, whereas the DP for (24) has a #P, with a singular number. Thus, whenever countable structure is introduced, either via NUMBER-agreement from a singular or plural demonstrative or plural morphology itself, the kind reading becomes unavailable. This suggests that the presence of #P, whether plural or singular, blocks kind reference. Then, definite kinds in English fail to project #P at all and are attributed the structure in (25), while definite individuals are structurally more complex (26).

(25) $[_{\text{DP}}$ the $[_{\text{NP}}$ N]

(26) $[_{\text{DP}}$ the/this/these $[_{\#\text{P}}$ {+sg/+pl} $[_{\text{NP}}$ N]]]

To provide support from the generics literature that there is no #P with countable morphology in the seemingly singular DS kind-referring subjects, let us consider more cross-linguistic data. If we expect #P to be present in definite individual structures, such as (26), then we should be able to infer information about the NUMBER of the N. However, seemingly singular predicates in Spanish (27) (Borik & Espinal, 2012, p. 129, their (13a) and Brazilian Portuguese (28) do not allow us to make inferences about their NUMBER:

(27) Este professor tiene libro.
This professor has book.
'This professor has a book/has published books.'

(28) Este professor tem aluno.
This professor has student.
'This professor has a student/students.'

These examples exemplify precisely this intuition: *The professor* in (27) can have one or more published books and the one in (28) one or more students. Therefore, an unmarked nominal, which might intuitively be categorised as singular, looks as if it is not marked for NUMBER at all. To extend the argument to kind-level predicates, they occur with the definite determiner and are not marked for plural, as in (29) (from Borik & Espinal, 2012, p. 129, their (14b):

(29) tener la gripe (porcina)
have the flu (swine)
'to have the (swine) flu'

Based on the theoretical assumption argued above, we can assume that if Spanish projects #P, as in (26), we would expect NUMBER-marking to be licensed. However, this is unacceptable in the plural (30) and cannot occur without the definite determiner (31) either (Borik & Espinal, 2012, p. 130, their (14c) and (15b)):

(30) *Tener las gripes porcinas
have the.PL flu.PL swine.PL
'to have the swine flu'

(31) *Gripe (porcina) está por todas partes
Flu (swine) is around all parts
'The (swine) flu is widespread.'

These examples underpin that a kind-referring subject in Spanish needs to be NUMBER-less. Further, bare nouns are not kind-referring in Spanish and cannot occur with kind-selecting predicates such as *to be widespread.* Note that no NUMBER is implied in these 'singular' Ns and that they cannot occur with plural morphology if they are kind-referring subjects. This provides further evidence for the hypothesis that NUMBER is responsible for predicating properties of kinds. If kind-referring DS subjects project neither singular nor plural, they must instead be NUMBER-less when referring to kinds directly, following Borik & Espinal (2012).

In addition to these observations, note that DS subjects of collective nouns, which typically allow plural agreement in British English (32-a), are restricted to 'singular' agreement in kind-referring predicates (32-b), due to the fact that when they occur as kind-referring subjects, they lack a Num-head.

(32) a. The committee was/were unhappy about the decision. [specific]
b. The committee was/*were invented to inhibit progress. [kind]

They surface as SG or PL because this ‘singular’ agreement is the default structure for agreement with numberless kind-referring DS subjects in English. Dayal (1999, p. 38, referring to Barker, 1992 and Schwarzschild, 1996 for further discussion) argues that “common nouns like team or committee ... are conceptually plural but gramatically [sic] singular and do not allow predication to individuals that we intuitively consider their instantiations.”

This supports the theory that kind-referring subjects are truly definite numberless (DN) subjects (although for the experiments below, I will continue to refer to them as DS subjects, since the contrast between kind- and non-kind-referring subjects is clear as no episodic reference is intended in any of the generic wording conditions). In English, we saw that plural morphology on the definite subject blocks kind readings (32-b) and only allows for individual readings, accounting for the difference that started this investigation, i.e. between *The tiger has stripes* and *The tigers have stripes*.

Consequently, all interpretations for nouns denoting instances of a kind, i.e. (groups of) individuals, and not kinds directly, require the intermediate morphosyntactic projection #P. This projection mediates between the semantics of the N-head, i.e. the root of N, denoting properties of kinds, and instances thereof. In terms of their denotation, the kind-referring DN differs from non-kind-referring DS with respect to its functional projections. #P is not projected in the kind-denotation in (33), since plural morphology is blocked, as is the absence of the definite determiner.

(33) a. $[_{DP}$ D $[_{NP}$ N$]]$
b. $[\![$D N $]\!] = \lambda x^k\ [P(x^k)]$
c. $<e^k>$

Analogously, the object denotation requires the presence of NUMBER in #P, as in (34-a), with the semantic denotation of Num in (34-b), the full DS in (34-c)) and the object denotation in (34-d) (adapted from Borik & Espinal, 2012, p. 131f, their (17) and (20):

(34) a. $[_{DP}$ D $[_{\#P}$ Num $[_{NP}$ N$]]]$

b. $[\![\text{Num}]\!] = \langle e^k, e^o \rangle$

c. $[\![\text{D Num N}]\!] = \lambda x^k\ \iota\ x^o\ [P(x^k) \wedge R(P(x^o), P(x^k))]$

d. $\langle e^o \rangle$

In (33), the definite determiner, represented by the iota-operator ι, applies to the kind-variable directly to denote properties of kinds, while the properties of kinds are in a realisation relation R to the properties of the objects picked out in (34-a), in line with theoretical assumptions of determiners in Section 2.2. By contrast, DS subjects with #P make reference to members of the kind, as ι cannot apply directly to the N to express kind reference.

Therefore, we can take kind meanings to originate in the N's root denotation. The insertion of #P prevents a kind reading for the subject, which in English co-occurs with singular or plural and in classifier languages with the insertion of a classifier, which likewise introduces NUMBER.

At the beginning of this section, I raised the example of Armenian, in which classifiers and plural-marking cannot co-occur. Taking a broader cross-linguistic scope, note that there exist languages in which classifiers and plural-morphology co-occur. This then means that they might not be in complementary distribution. Butler (2012, p. 22) raises the case of Yucatec Maya (contra Borer, 2005; Chierchia, 1998; Greenberg, 1963; Sanches & Slobin, 1973) with the examples (35)-(36) (taken from Blair & Vermont-Salas, 1967, p. 454, her (34)-(35)):

(35) ka'a-túul x-ch'úupal-o'ob
two-CL.AN FEM-girl-PL
'two girls'

(36) óox=p'éel siidra-o'ob-o'
three-CL.IN year-PL
'three years'

Butler further points the reader to Jakaltek Maya, in which two types of classifiers co-occur with plural morphology, illustrating that both numeral and noun classifiers do not stand in complementary distribution with plural marking (Butler, 2012, p. 23, her (37)-(38), taken from Craig, 1986, p. 15):

(37)

ca-waň	heb'	naj	winaj
2-CL.NM.HUMAN	PL.HUMAN	CL.N.MAN	man

'the two men'

(38)

ca-c'oň	(hej)	no7	nok'
2-CLN.NM.ANIMAL	PL.ANIMAL	CL.N.ANIMAL	animal

'the two animals'

These examples (among others from Halkomelen, see Wiltschko, 2008) do not support the hypothesis that the ClP and NUMBER-marking are in complimentary distribution. One way to account for this distribution would be that in some languages, classifiers and plural perform the same function and can both be used as evidence for NUMBER to instantiate individuals from a mass denotation. However, based on these data we could also stipulate different types of plural. Based on their position in the DP, they perform different functions, i.e. that "identical syntax (head of Number Phrase) does not follow from identity of function (pluralization)" (Butler, 2012, p. 44). Under such an interpretation, the co-occurrence is justified by the DP-adjoined position of the plural marking, unlike the English plural, which is further down in the DP, namely in the #P.[6] The main upshot is that the proposed analysis remains viable, albeit with the caveat that these Indo-European languages only have one type of plural and/or classifier.

Let us briefly look at these theories as these issue draw attention to a bigger question of whether the plural in a given language is analysed (or, following Chomskyan theories, merged) as a head or an adjunct within the layers of the DP. Both Wiltschko (2008) and Butler (2012) address this question in more detail and lay out diagnostics, such as optionality and agreement to determine which type of plural a language possesses. Looking at the diagnostic of optionality in plural marking, it is important to point out that the lack of plural marking allows for both singular and plural interpretations of a phrase can have alternative explanations. Butler (2012, p. 28) exemplifies this for Yucatec Maya with example (39):

[6] By extension, languages with both a #P and a ClP should theoretically allow the co-occurrence of plural marking and classifiers. An analysis of Armenian classifiers and plural marking is not part of the scope of this thesis, but should shed some light on this issue.

(39) le x-ch'úupal-o'
DEF FEM-girl-D2
'the girl (there)' / 'the girls (there)'

For the purposes of this thesis, it suffices to acknowledge that the co-occurrence of plural marking and classifiers in Yucatec Maya and Jakaltek Maya do not have the same effect as they would have in Armenian, presented in (20) above. I will not draw inferences on the position of the plural in English based on these examples, seeing that they likely do not have identity of form.

Under an analysis of generally *inclusive plural* as the default option and singular or *exclusive plural* readings as the exception (e.g. Arka & Dalrymple, 2016), this would not be a surprising interpretation. There are analogous examples in English, where the singular can be viewed as the marked case instead of the plural, assuming a separation between semantic and morphosyntactic markedness. This becomes obvious when referring to units of measurement, which are only singular for exactly n = 1. In both (40-a) and (41-a), the plural marker is necessary, despite the fact that the expressed amount is less than one:

(40) a. We have 0.8 mile*(s) to go.
b. We have 1 mile to go.
c. We have 2 miles to go.

(41) a. It is -0.5 degree*(s) Celsius outside.
b. It is 1 degree Celsius outside.
c. It is 20 degrees Celsius outside.

Note also that this is a stark departure from the NUMBER systems proposed in this thesis, which rely much more heavily on singular denoting a single instance, following Link (1983); Chierchia (1998). These cases would have to be re-interpreted and the most likely explanation under the framework proposed in this thesis would rely on the syntactic layers of the DP introducing individuation based on the plural marker in e.g. (40-a) or (41-a), as opposed to a distinct semantic interpretation of the singular in the corresponding examples (40-b) and (41-b).

Even though Butler (2012) argues that the absence of plural marking in Spanish implies singular reference, these examples hold in Spanish too[7]:

(42) a. Nos quedan 0.8 millas para llegar.
b. Nos quedan 1 milla para llegar.
c. Nos quedan 2 millas para llegar.

(43) a. Hace 0.5 grados afuera.
b. Hace 1 grado afuera.
c. Hace 20 grados afuera.

Furthermore, NUMBER agreement between determiner and noun is often raised as a marker of whether the plural is a head or an adjunct. For further discussion, I refer the reader to Brazilian Portuguese in which this agreement is not obligatory (an exception to this is written language of a formal register, due to diglossia in Brazil):

(44) a. As menina(s) gostam de engenheiria.
DET.FEM.PL girl.PL like.PL of engineering.
'The girls like engineering.'

The next section will take the findings about NUMBER and explore the interaction of DEFINITENESS and NUMBER.

2.3.1 DEFINITENESS and NUMBER

This section will take the observation of the complex relationship between a generic subject and its properties as a point of departure. With the observations about NUMBER from Section 2.3 in mind, I will explore how the morphosyntactic form of the subject can impose restrictions on its properties, i.e. how the observed linguistic limitations are justified.

Let us begin by looking at the three generic subjects for count nouns in English in turn. First, bare plural (BP) subjects, where semantically, the plural represents sets of (intensional) atoms, such that the generic BP is hypothesised to correspond to groups of individuals (Link, 1983; Krifka et al., 1995). Groups of atoms allow

[7] I thank my friend Irene País for translation and judgements in Spanish.

for averaging over properties, which corresponds to allowing properties that are merely statistically connected to the kind. I will present further details on the role of groups and averaging in Section 2.3.2.

Second, the indefinite singular (IS) denotes (intensional) atoms of individuals.[8] It does not represent groups and can therefore not perform any averaging over members of groups. The individual atoms are expected, under normal circumstances, to each possess the predicated property. This still leaves room for accidents, genetic mutations, etc. that lead to instances that do not possess a typical property of their kind. Overall, this corresponds to the observed distribution of principled properties occurring with IS subjects.

Third, the definite singular (DS) not only makes generic reference, but also allows for direct kind reference (following Borik & Espinal, 2012, 2015). When occurring with kind-level predicates, the properties do not only apply to instances of the kind, but to the kind itself. For example, when asserting generic statements such as *The woolly mammoth went extinct after the last Ice Age*, no statement is made about a specific woolly mammoth, and neither about a majority group of normal mammoths, since only a kind can go extinct.

To elaborate on the notion of kind-level predicates and the role of DS subjects, let us take a kind as an abstract entity that ties together stages and individuals. It can occur with kind-denoting predicates such as *be invented* or *be extinct*, as first introduced by Carlson (1980). These kind-denoting predicates cannot apply to individuals since no individual can be extinct, only the kind (Dahl, 1975; Carlson & Pelletier, 1995).

It is certainly not the case that the BP, just because it is the most ubiquitous generic subject form, is always the preferred or most appropriate one. Take, for example, the generic statement (45):

[8] This also leaves room for speculation about IS DPs in non-subject positions. While cooking a curry one night I was repeatedly reminded by my friend that they have "never eaten a mung bean" and that they do not know where "a mung bean stands in [their] bean hierarchy". Clearly, these statements did not make reference to a single mung bean (which I confirmed as I was delighted to spot the usage of an IS generic in the wild).

(45) a. The dodo is extinct.
b. *A dodo is extinct.
c. ?Dodos are extinct.

In cases with kind-level predicates, the DS subject is typically the most appropriate subject form. In fact, the BP subject is only felicitous in certain circumstances. In these cases, the acceptability is determined by the type of predicate the generalisation co-occurs with. Kind-level predicates, such as *be extinct*, cannot apply to individuals and typically do not co-occur with groups. Direct kind reference, achieved via the DS subject, is the best option. Mueller-Reichau (2013, Chapter 5.5) investigates these in detail and further categorises them into subgroups of kind-level predicates. These judgements are commonly accepted and the reason why the DS is often seen as the subject form that best applies to kind-level predicates. However, Mueller-Reichau brings to attention the following example (2013, Chapter 5.5, his (17)):

(46) a. A trullala has been invented.
b. *A trullala is widespread.

which clearly indicates that an IS subject can co-occur with a kind-level predicate. Arguments that the predicate *be invented* is not truly a kind-level predicate do not hold for the following reasons. In very broad terms, no individual can be invented, only a kind can. Before providing a possible explanation of this, I will introduce a related approach to the semantics of the generic subjects provided by Prasada.

Prasada (p.c.) comes to similar conclusions about the tripartite setting of syntactic and semantic interpretations, albeit from a psychology perspective. He focuses on characterising kinds and argues that there are three distinct mechanisms of characterising kinds, each of which is fulfilled by one of the three generic subjects in English. The first way of characterising kinds, he argues, is via “characteristics of a sample of the kind”, in other words by using an arbitrary sample. This is conveyed by the BP in English: The arbitrariness is provided by the indefiniteness of the subject, the aspect of it being a sample by the plural number of the subject. The second way is by “an arbitrary instance of the kind”. Following the same logic as in the first way, an arbitrary instance is provided by an indefinite subject. As this way characterises kinds not via a sample, but an instance, this subject is singular. The IS is therefore expected, and found, to be most appropriate for generic statements about kinds that do “not involve information that is specific to the arbitrary instance”, such as the sum of the internal angles of a triangle. The third way Prasada proposes is via exemplification, that is “by interpreting a specific instance of the kind as exemplifying the kind”. Note that this way of characterising kinds requires a specific instance, therefore the subject must be definite, and it requires an instance, thus the English DS is the appropriate subject for this type of characterisation. Prasada further excludes superordinate kinds from this characterisation and only allows for basic level kinds to be used, contrasting *The parabola is easy to draw* with *#The curve is easy to draw.*[9] Prasada comments that this account predicts PCs, but not SCs, to be acceptable given that uniformity in the kind-characteristics connections is assumed. Principally connected properties are normally expected to be uniformly present for the kind, whereas merely statistically connected properties are subject to more variation. Similarly, we expect more variation to be present in social kinds as opposed to mathematical objects, which explains how the DS expressing PCs is not as acceptable for groups of people, e.g. *#The Canadian is polite*, as it is for mathematical objects, e.g. *The triangle is three-sided.*

Now, let us return to the issue raised in (46-a), and how an IS subject can easily co-occur with a kind-level predicate. These cases can be accounted for if this *trullala* stands in as a sample for the kind. However, some seemingly parallel

[9]However, this raises the question of category-property links and their influence on acceptable generic subjects once more. If the judgements for these two sentences favour a subkind over a basic level kind, then the standard example of “The triangle is three-sided” becomes more difficult to account for. These examples of mathematical objects have been discussed in Carlson (2010, p. 30), with reference to Vendler (1971):

(i) a. *The parabola* is easily plotted.
b. ?? *The curve* is easily plotted.

Note that to this, a reviewer raised further examples of *the curve* being acceptable in other contexts (to varying degrees of acceptability), which at the very least shows that casting a general net of categories being ‘too general’ is not straight-forward. I would argue that the felicity of these judgements is dependent on the specific category-property links targeted in each instance.

examples do not allow the same subject-predicate combination, complicating the pattern, see (47) (adapted from Krifka et al., 1995, p. 12):

(47) *A rat was introduced to Australia in 1770.

Instead, one might argue for a subkind interpretation that becomes available via the IS. For (46-a), this implies the previous existence of 'the trullala', and the most intuitive interpretation of *a trullala* in this context to be that 'One (kind of) trullala' has been invented. This contrasts with (47), however similarly, one could coerce a subkind reading here, arguing that 'One rat, namely the dusky rat, was introduced...' allows for non-DS subjects with kind-level predicate.

Recalling the arguments for NUMBER-less kind reference in examples (54)-(56) above, I will now join together the pieces of this discussion by proposing a linking hypothesis between the psychological and the linguistic side of genericity, via the presence or absence of the NUMBER projection.

2.3.2 Groups and averaging

Once individual members of a kind have been instantiated, we must further allow them to have properties as a group. As generics are known to be exception-tolerant, statements about the kind may be judged as true, even in the face of individuals lacking the predicated property. One way of accounting for a generally true statement of a group, without requiring all instances thereof to possess the property, is averaging over instances. To do so, we may stipulate semi-lattice structures as proposed by Link (1983).

Plural individuals are sets of singular individuals, which can be represented in a join (or sum) semi-lattice structure (adapted from Chierchia, 1998, p. 345, his (9a), see Figure 2.1). This idea was first brought forward by Link (1983) who used lattice structures to account for the semantics of singular and plural, where individual letters (*a, b*, etc.) stand for atomic individuals, which we assume to be present in singular subjects. Groups of individuals are joined in curly brackets

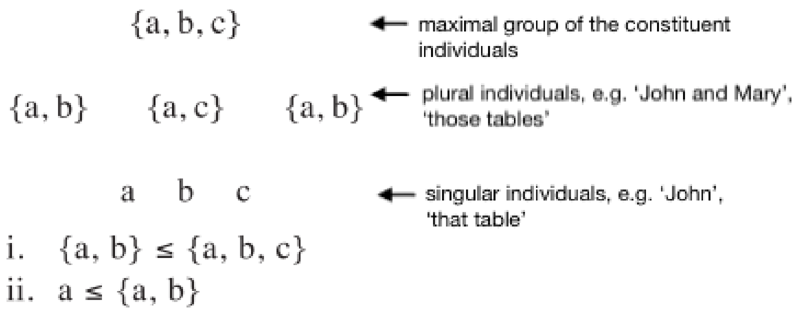


Figure 2.1: Chierchia (1998)'s join semi-lattice structure.

Figure 2.2: Carlson (1980)'s semi-lattice structure for s-, o-, and k-level individuals.

(*{a, b}*, etc.), and the maximal group is the conjunction of all atomic individuals (denoted by ι in definite descriptions).

In Chierchia's framework, individuals are turned into sets via pluralisation, where two atomic individuals are joined by the join operation $\cup$ such that the operations are equivalent as in (48).[10]

(48) a $\cup$ b = {a, b}
= smallest set of pluralities containing atomic individuals a, b

By adapting the framework of lattices, introducing a plural nominal means that atomic instances, i.e. individuals denoted by the corresponding singular nominal, are included in the denotation. Similarly, a join semi-lattice structure can illustrate the levels of stage-, object-, and kind-level individuals (Carlson, 1980, p. 69, see Figure 2.2).[11]

Why should these semi-lattice structures be able to account for the felicitous uses of BP generics? While these illustrations of groups are useful to interpret

[10]It remains unclear how exactly Chierchia accounts formally for the retrieval of the mass term under the join operator, though it might be along the lines of denoting a singleton set: $\cup$ a = {a}.

[11]Note that Carlson thereby describes the notion of stages as primitive and that kinds serve to tie both objects and stages together, as they can consist of several stages or various objects (1980, p. 69).

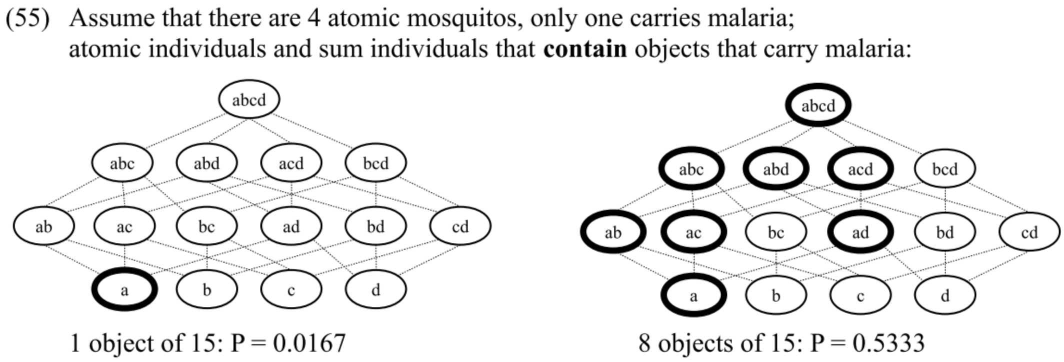


Figure 2.3: Krifka (2017)'s semi-lattice structure for striking properties.

features of groups of individuals and objects, we have already seen that generic statements do not follow the normal subset relationships of quantified statements. To address this question, Krifka (2017) argues that (plural) generic statements are in fact about samples and not about a whole set, mirroring Prasada's (p.c.) argument. Since we would expect large variation in any sample, in order for the statements to be appropriate, Krifka (2017, p. 3) continues that "samples are chosen at random" in order for them to be useful, and further that "[t]he information that generic sentences convey is (...) about what one should expect when encountering elements of the restrictor set."

With such samples, one would expect majority characteristics and statistically prevalent properties to be expressed easily, as we find in the examples. Yet, we have also seen that BP generics, the only plural generic subject in English, is also the only generic subject that may express striking properties. Under Krifka (2017)'s account, these are straightforwardly accounted for. Only a small percentage of individuals need posses the property in order for the whole group, represented as the sum or join individuals, to possess them in over half of the cases, due to the spread of properties through the joined structures. Krifka (2017) uses the example of *Mosquitoes carry malaria* and represents it visually (Figure 2.3).

This is a departure from Borik & Espinal's proposal, who argue that kind reference is due to the absence of NUMBER in the DP, and therefore a hallmark of the DN. Prasada does not propose a syntactic analysis of generic noun phrases but instead relies on evidence in the psychology literature, drawing on examples such as the DS being the best generic subject for basic kind-level reference. These approaches could be unified if we maintain the special NUMBER-less status of DS subjects and allow for #P to be present in IS and BP subjects. Then, any presumed instantiations can still follow the semi-lattice structures proposed by Link, Carlson, and Krifka, as introduced in this section.

2.3.3 Reference to subkinds

The final component in this account of the role of NUMBER is a theory of subkinds. In terms of terminology and definitions, there is a certain level of ambiguity, and even subjectivity, when defining the notion of a 'subkind' against basic level kinds and superordinate kinds. These terms "can be considered as partial orderings or lattice structures" (Krifka et al., 1995, p. 76). Let me begin by making explicit my assumptions about basic level kinds. Krifka et al. (1995, ibid) have defined them as "frequent and noncomposite (...). The words at this distinguished level are most readily used to describe an object". By extension, subkinds convey less frequent and composite meanings. They pick out a level in the kind hierarchy that is below the basic kind level, mirroring the process for picking out a superordinate kind, above the basic kind level. Further discussion about the acquisition of categories as well as categorisation of instances from a more psychological-conceptual perspective is provided in Chapter 6. In this section, I will focus more specifically on some linguistic observations and theories that aim to provide an isomorphic account of syntactic structures and semantic interpretations. Note also that this discussion will not encompass the creation of novel concepts through linguistic innovations, such as nonce-uses and neologisms, although this has been discussed in the literature with examples such as *The student teapotted the dean* and minimal pairs like *The boy porched the newspaper* and *The builder porched the house* (presented in Murphy & Medin, 1985, with reference to Clark & Clark, 1979).

The previous sections highlighted the close relationship between NUMBER and realisation by contrasting kind- and object-referring DPs. This section further investigates the role of these morphosyntactic projections by moving away from contrasting kind reference and object reference, and instead constrasting kind reference and subkind reference. Given that kind-referring subjects are taken to be NUMBER-less, how is reference to subkinds possible with plural morphology, which is argued to originate in #P? Consider (49)-(50), based on Zamparelli (2000). Once NUMBER is inserted, the kind-reading becomes unavailable and the subject is instead analysed as a subkind. This also holds for mass-nouns (51):

(49) a. The whale is almost extinct. [kind]
b. The (two) whales are almost extinct. [subkind]

(50) a. *The kind of whale is almost extinct. [kind]
b. This (one) kind of whale is almost extinct. [subkind]
c. These (two) kinds of whale are almost extinct. [subkind]

(51) a. Juice contains a lot of sugar. [kind]
b. One juice, specifically green juice, is healthy. [subkind]

These example reiterate that kinds cannot occur in the plural. When the same statement has a plural subject, reference to subkinds is made, which do not resist NUMBER-marking (49). They further permit the insertion of *kind of*, which requires some expression of NUMBER (50), and even mass terms become countable (51).

There is similar variation with the insertion of the quantifier *some*. In the context of psychological research this variability has been pointed out by Gelman et al. (2010, p. 9, on the role of world knowledge):

(52) a. I like rice. [kind]
b. I like some rice. [subkind]

The shift from a kind reading to a subkind reading is analogous to (49)-(51), as the insertion of *some* shifts (52-a) to a subkind reading (52-b). Semantically, this provides further evidence that #P realises properties of individuals, either via a classifier or via true plural marking, and shifts a semantic entity from the domain of kinds into the domain of objects (Link, 1983). If kinds and subkinds differ syntactically with respect to NUMBER, then the structure proposed for object denotation also applies to subkind interpretations, repeated from (21) (p. 26 above):

(53) $[_{DP}$ D $[_{\#P}$ Num $[_{NP}$ N$]]]$

This continues to allow an interpretation of truly kind-referring expressions to be NUMBER-less, but subkinds are built on NUMBER. This also explains why subkinds do not resist NUMBER-marking and syntactically project #P. Accepting the notion that whenever #P is present, the properties of the kind denoted by N are realised through the shift to properties of individuals, implies that spatiotemporal boundedness is linked to the realisation of properties of (groups of) individuals. Under such an analysis, subkind interpretations are built on NUMBER and always spatiotemporally bounded.

Let us return to Borik & Espinal's analysis of kind reference. Recall that they depart from the assumption that NUMBER may take one of two values, singular or plural (such an assumption might be assumed implicitly, as in Gelman, 2004). Instead Borik & Espinal (2012, 2015) assume that there are three distinct 'values' in English: singular, plural, and numberless (for IS, BP, and DS, reanalysed as DN, respectively). DN subjects with kind-referring predicates are argued to be NUMBER-less and not singular because they resist NUMBER marking and do not permit the insertion of *kind of* without simultaneous expression of NUMBER. Support comes, among others, from Spanish, where kind-referring subjects are grammatical only if no overt NUMBER expression or other lexical items are inserted (54)-(56) (adapted from Borik & Espinal, 2012, p. 133-135):

(54) La nevera se inventó en el siglo XVIII.
The.SG fridge CL invented in the century XVIII.
'The fridge was invented in the 19th century.'

(55) *Las (dos) neveras se inventaron en el siglo XVIII.
The (two) fridges CL invented in the century XVIII.
'The (two) fridges were invented in the 19th century.'

(56) *La clase de nevera se inventó en el siglo XVIII.
The class of fridge CL invented in the century XVIII.

‘The type of fridge was invented in the 19$^{\text{th}}$ century.’

With these examples, I maintain a linking hypothesis between the psychological and the linguistic side of genericity via the projection of NUMBER. Recall that semantically, numberless nominals range over the domain of kinds, which refer to intensional objects, whereas group-readings refer to plural entities, which are numbered or ‘instantiated’ nominals in the domain of objects (e.g. Link, 1983). Contrasting with DNs, the presence of NUMBER in IS and BP generic subjects is related to realisation, not licensed in kind-referring subjects, and thus there is no NUMBER marking in truly kind-referring subjects[12].

Following these observations, in Kwapiszewski & Fuellenbach (2020) we pursue the theory put forward by Borik & Espinal that the presence of NUMBER is responsible for non-generic, object-level interpretation of subjects that can otherwise be interpreted as generic, or more specifically kind-referring. By presenting the following data on Polish kinds and subkinds, we provide further evidence that subkind reference is also reliant on NUMBER. It is thereby categorically different than kind reference, which remains NUMBER-less (Kwapiszewski & Fuellenbach, 2020):

(57) Jeden tygrys jest na skraju wymarcia.
one tiger.NOM is on verge extinction.GEN
‘One (kind of) tiger is on the verge of extinction.’

Example (57) illustrates that the insertion of a lexical marker, crucially a NUMBER-ed one, such as *one*, allows for a shift to subkinds. One alternative way to achieve subkind reference (which will become more relevant in Section 2.5) is via kind modification. To exemplify, for the NP *Bengal tiger*, the kind modifier *Bengal* selects a specific subkind (or set of subkinds) from the denotation of *tiger*. Example (58) features the classifying adjective *bengalski* in Polish.

(58) Tygrys bengalski jest na skraju wymarcia.
tiger.NOM Bengal.MASC is on verge extinction.GEN
‘The Bengal tiger is on the verge of extinction.’

[12]For further evidence, see Borik & Espinal (2012) on Russian, or Fuellenbach (2017, unpublished) for a wider overview.

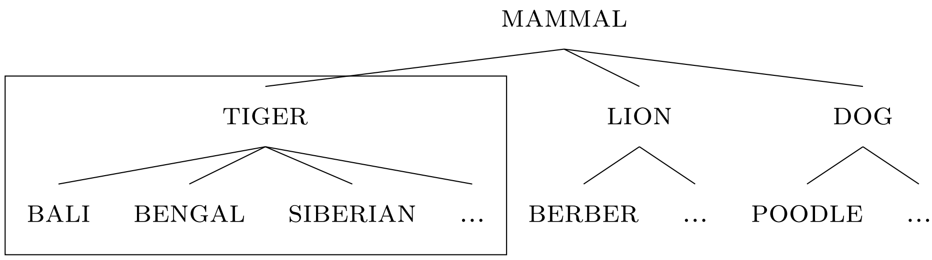


Figure 2.4: Taxonomic NP Denotations

In order to account for these types of subkind reference, in Kwapiszewski & Fuellenbach (2020) we refer to work from McNally & Boleda (2004) on relational nouns in Catalan, as well as to Wągiel (2014) on classifying adjectives in Polish, and Borik & Espinal (2015) on kind modifiers in Spanish. Following their approach, the composition of nouns and their modifiers is intersective, proceeding via the composition rule of *predicate modification* (see Heim & Kratzer, 1998). With regard to (58), this leads to a set of kinds, which can be denoted by $[\![\text{tiger}]\!] = \{$BENGAL TIGER, SIBERIAN TIGER, ... $\}$, which intersects with the set of kinds denoted by $[\![\text{Bengal}]\!] = \{$BENGAL TIGER, BENGAL CAT, ...$\}$, yielding the correct denotation for the modified NP.

Assuming that kind reference in Polish is provided via the ι operator, as it is in English, let us continue to investigate how intersective kind modification under previous accounts is troublesome as "the fact that the ι operator can apply to $[\![\text{tiger}]\!]$ in [the semantics of Bengal tiger] entails that $[\![\text{tiger}]\!]$ is a singleton set containing only the maximal kind TIGER. In other words, this derivation assumes that NP denotations are atomic" (Kwapiszewski & Fuellenbach, 2020, Section 4, *Incompatibility with Intersective Kind Modification*). The issue is that this does not allow for intersection with any modification, be it *Bengal* to render *Bengal tiger*, or of *Bengal* with any other kind, e.g. to render *Bengal cat.* Instead, we propose that subkind reference is achieved by ascribing taxonomic denotations to NPs, which are introduced via NUMBER or kind modifiers. This allows to maintain NUMBER-less direct kind reference, in line with Borik & Espinal. This denotation is represented graphically in 2.4, as in Kwapiszewski & Fuellenbach (2020).

Ultimately, we adopt a theory that maintains an atomic semantics for basic kinds and uses function application of modifiers to derive a subkind. We propose that such a function is a subkind operator, or more specifically a “dyadic relation between kinds and their subkinds” (Kwapiszewski & Fuellenbach, 2020, Section 4, *Towards a Solution*). This idea is similar to Krifka et al. (1995, p. 77)’s taxonomic subkind relation *T*, although Krifka’s focus is on the relationship between the domain of kinds and the domain of objects, whereas this subkind operator specifically relates atomic kinds to taxonomic subkinds.

The main point of the proposal presented in Kwapiszewski & Fuellenbach (2020) is that a SubkindP, as a type of ClP, is proposed to account for intersection as well as correct linearisation of sentential constituents. ClP can either introduce a realisation operator, which allows for object interpretations, or it can introduce the subkind operator, which allows for subkind interpretations (as both of these rely on NUMBER to be present through #P). This is still in line with the proposal of definite, NUMBER-less kinds, and independently observed in Picallo (2006), who argues that the projection of #P is dependent on the presence of ClP. Overall, we adapted Borik & Espinal (2012, 2015)’s proposal of the structure of kinds (and by extension an explanation of the co-occurrence of NUMBER and subkinds), considering Borer (2005)’s covert layers in the DP, and Kratzer & Selkirk (2007)’s function of ClPs.

Integrated in the extended DP structure proposed above, a subkind nominal is expected to have the structure in Figure 2.5, as in Kwapiszewski & Fuellenbach (2020).

Due to the presence of #P, and the classifier, no direct kind reference is possible. This structure therefore shows the more easily interchangeable object-level and subkind-level references. Most importantly, such a structure continues to support the close link between syntax and semantics in the DP.[13]

[13]Further research into the semantics and cognitive make-up of subkinds should take into account how kinds share subkinds. For instance, the iPhone is a subkind of both the phone and the computer (or even the calculator). This implies that kinds need to be able to share subkinds. A related idea would be to consider the implications of manipulation the number of properties that are shared at the kind level. Instances of the kind PAPERCLIP should share almost all (visual) properties whereas the kind DOG displays huge variation. An interesting datapoint would be the level of rejection of kind membership when not enough properties are shared. For further discussion on instances of kinds, see Prasada & Hall (2019). Another related area of research concerns the diachronic development of subkinds and overall kind. For example, carrots, which

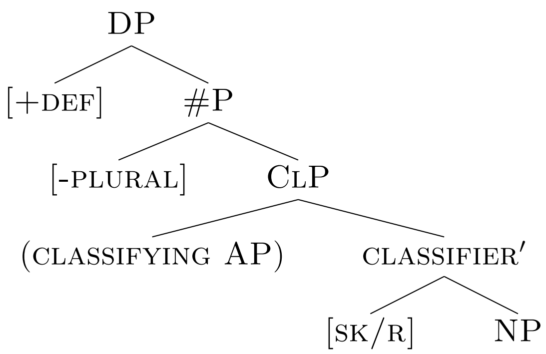


Figure 2.5: The Structure of a Definite (Modified) Subkind Nominal, in which #P requires ClP.
NB: The indicated position for R indicates the corresponding structure for an object-level nominal

This section has addressed the role of NUMBER for the three generic subjects for count nouns in English. I argued that in this respect, DS subjects are distinct in their morphosyntactic layers by not projecting a #P and thereby referring to kinds directly, following Borik & Espinal (2012, 2015)'s analysis. Both IS and BP subjects project #P; the former with a singular value and enabling subkind interpretations, the latter with a plural value and introducing join semi-lattice structures that allow for averaging and statistical properties of kinds.

Crucially, this section has argued that there exist a variety of structures for generic subjects in English. The initial hypothesis put forward in this thesis is that these will be used in different contexts, in line with their distinct semantics. Further, the prediction for the experimental chapters is that these will results in distinct behavioural results. In particular, the experiments will test whether these distinct morphosyntactic forms effect distinct behavioural responses, such that some generic subject types are better suited to support certain types of properties predicated of the kind.

are typically orange and straight, use to be purple and curly. This question is intertwined with questions about actual existing vs possibly existing kinds, both in the past and the future. To wit, what if these curly purple carrots were actually ‘shmarrots’ which slowly developed to carrots, and will in the future turn into yellow, rotund vegetables, called ‘barrots’? It might even be that shmarrots and barrots would have the same nutritional profile. We must ask ourselves if these would be classified as distinct (sub-)kinds. A change of principled properties, occurring slowly over time, raises questions of continuous identity. For further discussion, see also the metaphysics problem of the Ship of Theseus, discussed by Heraclitus and Plato.

The next section will explore two distinct perspectives that have been proposed for NUMBER-ed generic subjects: *Generics-as-Default* and the generic operator GEN. I will compare these approaches, and focus in particular on GEN, which has been proposed to account for the ambiguity between generic and individual interpretations of (singular) generic subjects.

2.4 *Generics-as-Default* or GEN?

Having discussed the range of subjects that allow for generic interpretation from a morphosyntactic point of view, the question still stands as to what distinguishes the NUMBER-ed generic subjects from their episodic counterparts. I have addressed the role of DS subjects (for direct kind reference, mediated by the absence of #P), the role of the BP (accounting for groups and averaging), and the role of IS (analysed as making reference to an arbitrary individual that exemplifies the kind, following Prasada, 2016a). The main discussion of this section presents accounts of generic subjects with and without the generic quantifier GEN (in Sections 2.4.1 and 2.4.2, respectively), and an outline of IS and BP subjects with GEN.

I will begin by presenting some approaches that have stipulated generics as a default for the conceptual system, providing an overview of the generic overgeneralisation effect (GOG) and its relationship to *Generics-as-Default.*

2.4.1 Generics without GEN

One proposal addressing the lack of a generic subject's direct correspondence to a quantifier is the "Generic Overgeneralisation Effect". GOG was first explicitly mentioned in Leslie et al. (2011), in which generic statements are contrasted with quantified ones. Leslie et al. start with the observation that while generic expressions tolerate exceptions, universally quantified ones do not. In the studies, participants were presented with either generic (e.g. *Ducks lay eggs*), existentially quantified (e.g. *Some ducks lay eggs*) or universally quantified statements (e.g. *All ducks lay eggs*). When participants were asked to provide truth-value judgements [TVJs] on such statements, the results show that even when explicitly mentioning the fact that only healthy, female, and mature ducks lay eggs, or when presented with alternative statements (*Some ducks do not lay eggs*), participants judged *all*-quantified statements as true. In other words, participants answered questions like *Do all ducks lay eggs?* positively. The willingness to accept an *all*-quantified statement as true in the face of exceptions, or "the tendency to judge universally quantified statements to be true when the corresponding generic is true" (Leslie et al., 2011, p. 15) is the expression of GOG.

Leslie et al. argue to have ruled out that participants used either a subkind reading or domain restriction, by giving population information such as the total number of ducks, or that they ignored relevant facts, by providing correct alternatives (e.g. *Only some ducks lay eggs*; *Some ducks don't lay eggs*). Consequently, they take GOG to be the most likely explanation. While they also find it to be robust across different types of generics (2011, p. 14), they qualify that their examples mainly consisted of (minority) characteristic predications, since "these universals offered the clearest illustration of the Generic Overgeneralization effect" (ibid). At first glance, this offers a straightforward explanation and does not conflict with previous results from related work.

However, this conclusion is, once again, a result of taking genericity as a homogeneous phenomenon that can uniformly be pitted against quantified statements. In general, psychology is not unaware of the different conceptual types of generic expressions: on a theoretical level, (Leslie et al., 2011, p. 5) identify five sub-categories of generic statements:

1. quasi-definitional (*Triangles have three sides*)
2. majority characteristic (*Tigers have stripes*)
3. minority characteristic (*Lions have manes*)
4. majority statistical (*Cars have radios*)
5. striking generics (*Pitbulls maul children*)

Yet, the most popular subject type in generic experiments remains the BP, despite the fact, or possibly because of the fact, that it is licensed to occur with all sub-categories of generic statements. Psychology thereby does not take seriously the importance of the morphosyntactic variation of generic subjects based on linguistic factors.

The most prominent analysis of generics in psychology takes generics are the cognitive default. *Generics-as-Default* argues that a sentence, unless otherwise quantified, receives a generic interpretation without the need for additional covert quantification, based on the assumption that generic expressions are the default manner of reference (Leslie, 2008; Gelman, 2004, among others). Simply put, *Generics-as-Default* means that generic expressions are innate, acquired before quantified statements from which they are categorically different, and unmarked when compared to specific reference (*Birds lay eggs* vs *This bird lays eggs*) and quantified expressions (*Tigers have stripes* vs *All tigers have stripes*). *Generics-as-Default* proposes that the presence of an overt quantifier (e.g. *most, some, all*) transparently signals an additional process compared to the 'bare' generic expression (Leslie et al., 2011, p. 2). This proposal suggests that both from a linguistic and a psychological perspective, the form and standard interpretation of generics reflect the simplicity they are assigned as the cognitive default. Support comes from observations in acquisition (children acquire and use generic expressions earlier than quantified ones, e.g. Gelman, 2003, 2004) and targeted experiments (showing shorter response times for generic phrases as evidence for their cognitive ease, e.g. Meyer et al., 2011).

Generics-as-Default sounds promising: it accounts for ease of processing by aligning linguistic simplicity with cognitive simplicity. However, recall that linguists are primarily concerned with a formal account of genericity, which is more challenging than providing a set-theoretic formalism for quantified statements (Lazaridou-Chatzigoga & Stockall, 2013). Semanticists have worked on capturing genericity formally since at least the 1970s (Lawler, 1973; Dahl, 1975; Carlson, 1980) and need to respond to the surging interest in generics in psychology and the cognitive sciences. In this attempt, linguists have noted problems in the experimental set-up and methodological assumptions (see Lazaridou-Chatzigoga & Stockall, 2013;

Lazaridou-Chatzigoga et al., 2015, and discussion below). In joint support of GOG and *Generics-as-Default*, Leslie et al. (2011) discuss that GOG is potentially only restricted to the quantifier *all*. This raises the point that there might be more fine-grained distinctions within quantifiers, which is an intriguing parallel to the main argument of this thesis, i.e. that there are fine-grained distinctions within generics (see also Lazaridou-Chatzigoga & Stockall, 2013, p. 487).

In general, stipulating separate mechanisms for generic and episodic statements, in the form of *Generics-as-Default*, is not desirable. It is preferable to treat genericity-specific phenomena under an analysis that also covers other linguistic phenomena, such as domain restriction proposed by Lazaridou-Chatzigoga & Stockall (2013). They argue that, more importantly, not only did Leslie et al. not rule out their three alternative explanations, they also did not provide experimental evidence in favour of *Generics-as-Default* where processing of generics would be faster than that of quantified statements (2011, p. 487f). Lazaridou-Chatzigoga & Stockall (2013) address these proposed alternative explanations by carrying out timed TVJs, as was the case in the studies run by Leslie et al. (2011). In this version, participants were first introduced to a context, and then had to judge a statement as true or false. The subject of this statement was systematically manipulated and could occur in one of four conditions: generic, *all*, *all the* or *every* (e.g. {Tigers/All tigers/All the tigers/Every tiger} {have/has} stripes). Further controlling for the type of connection between subject and predicate (minority, e.g. *Ducks lay eggs*, vs majority characteristic, e.g. *Tigers have stripes*), the results suggest that while some generics might be easier to process, reading and reaction time for TVJs is dependent both on context and the type of generic and quantifier (Lazaridou-Chatzigoga et al., 2015, p. 488).

Finally, one of the more prominent arguments in favour of *Generics-as-Default* does not hold cross-linguistically: generics are not necessarily the least marked option to express knowledge about the world. Many Romance languages use definite plurals (e.g. French) which are marked overtly with both a determiner and plural morphology. In Greek, quantified and generic expressions are both overtly marked.

The former is usually expressed with a quantifier and a singular NP and the latter is most prominently a definite plural (see Lazaridou-Chatzigoga et al., 2015, p. 488).

In summary, greater variation in terms of the types of generic and quantified subjects is necessary to assess the validity of *Generics-as-Default* and the related GOG. In a similar vein, Lazaridou-Chatzigoga et al. (2015, p. 484) note a potential task effect, given that psychology experiments mainly use TVJs and recall-tasks with generic, universally quantified (*all*), and existentially quantified (*some*) sentences (2015, p. 484), ignoring other means of accessing the acceptability of various generic statements. Therefore, domain restriction is not only a possible, but potentially even a preferable explanation because it is a more general mechanism that can account for both *Generics-as-Default* and other linguistic phenomena. If domain restriction can account for our interpretation of generic statements, we need not stipulate a separate mechanism specifically for this task. Granted these criticisms, it remains to be seen how formal linguistics can account for the ease of learnability of generics, possibly with a covert modal operator, as discussed in the next section.

2.4.2 Generics with GEN

The fact that any generic subject can be interpreted as non-generic under different circumstances has given much room for debate on how these seemingly identical subjects differ. To formally capture the meaning of generic subjects, some semanticists analyse genericity as a case of quantification where the quantifier GEN binds the sentence's variables. Seeing that there is no overt quantification, compared to sentences such as *Most tigers have stripes*, this possible quantifier has to be covert, similar to #P and ClP (and SubkindP).

Krifka et al. (1995) introduces GEN as a dyadic operator with a restrictor and a nuclear scope, based on Carlson (1989). Example (59) is the reformulated and updated denotation in Krifka (2017, Section 2, Paragraph (3)):

(59) Typhoons are dangerous.

GENx (λx[typhoon(x)], λx[dangerous(x)])

This phonologically null quantifier resembles other adverbs of quantification, holds at least over cases (i.e. tuples of persons and a time-/event-coordinate, Lewis, 1975, p. 182), and is unselective with regard to the variables it binds. GEN quantifies over the entire sentence and binds any variables not bound by other quantifiers.[14]

This approach sheds light on a related issue: Based on the structure we have identified for DPs, how can a generic reading be achieved for sentences such as *A tiger has stripes*, but crucially not for *This tiger has stripes*? Let us take stock of the proposed structure so far: The indefinite determiner differs from the definite *the* as well as the demonstrative *this*. The indefinite determiner is subject to the same concerns given that the same structure with an IS subject can have an existential and a generic reading. I have argued that the indefinite determiner *a* fulfills the function of introducing divisibility (in ClP), and quantity (in #P), such that the structure with an existential reading can be argued for via existential closure. Existential closure applies whenever variables are within the scope of the existential quantifier ∃. When variables are outside the nuclear scope of ∃ (often equated to the c-command domain of the verb), ∃ cannot apply (Diesing, 1992).

Extending this logic, generic readings can be achieved via the proposed generic operator GEN. Under this analysis, GEN is as an unselective, covert quantifier, similar to overt quantifiers such as *usually* or *generally*, which can bind multiple variables, and is only licensed when existential closure (via ∃) is not available.[15]

To illustrate the application of GEN, consider the generic statement (60) which receives the denotation in (61) following Carlson (1980), using Heim (1982)'s framework:

(60) Brussel sprouts are unsuitable for eating.

(61) GEN x [x is a Brussel sprout] x is unsuitable for eating

[14] Carlson (1977, p. 43) introduces this generic operator as *G*. This G "maps predicates that apply to stages to predicates that apply to individuals." Formally, Carlson represents G as «s, <e^s, t»,<e^i,t».

[15] I will leave aside the issue that the application of such an operator is problematic in cases of striking generics where the property specifically applies to unusual instances of the kind (see e.g. Liebesman, 2011).

In cases of non-generic reference, these variables fall under existential closure as illustrated by Carlson (1980)'s parallel examples (62)-(63):

(62) Carpenter ants destroyed my viola da gamba.

(63) $\exists$ x [x is a carpenter ant $\wedge$ x destroyed my viola da gamba]

GEN is assumed to give rise to an interpretation of 'normality' as the modal base is taken as possible worlds that are as similar as possible to the actual world (e.g. Nickel, 2016, though generics cannot be contextually restricted, see von Fintel, 1996, p. 12f). The crucial aspect is that without existential closure, GEN binds intensionalised individuals. This is what allows generic reference, where generic subjects can refer to future or past individuals, or even individuals in counterfactual situations. Nickel summarises this modal aspect of generics by stating that they "seem to not just be substantially independent of statistical facts in the world and at the time with respect to which they are evaluated. They seem to have implications for worlds or times beyond" (2016, p. 16).

In German, the constituent structure of generics allows us to differentiate between existential and generic readings in subclauses with sentential particles via the relative position of the subject. Focusing not exclusively on subjects but the clausal structure, note that if the subject is not in [Spec, IP] but lower down in [Spec, VP], this is reflected in the surface structure by the subject being preceded by modal particles such as *ja doch* (roughly "indeed"). In this position, it can be existentially closed (64-a). On the other hand, if the subject precedes the modal particles, it is outside of the scope of the existential quantifier and therefore bound by GEN (64-b) (Diesing, 1992, p. 36, her (32)):

(64) a. ... weil ja doch Linguisten Kammermusik spielen.
... since PRT PRT linguists chamber.music play.
'... since there are linguists playing chamber music.' [existential]

b. ... weil Linguisten ja doch Kammermusik spielen.
... since linguists PRT PRT chamber.music play.
'... since linguists (in general) play chamber music.' [generic]

Note that these examples receive the same judgements with IS subjects. Since the indefinite determiner is not in D, <e> in DP is also bound by GEN.

Based on Webelhuth (1989) (inter alia for other languages), arguing that sentential particles mark VP boundaries in German, I adopt the following syntactic structure for existential (65-a) and generic (65-b) readings (based on Diesing, 1992, p. 31):

(65) a. ... [$_{\text{CP}}$weil [$_{\text{IP}}$ja doch [$_{\text{VP}}$Linguisten Kammermusik spielen]]].
... since PRT PRT linguists chamber.music play.
'... since there are (indeed) linguists playing chamber music.'

b. ... [$_{\text{CP}}$weil [$_{\text{IP}}$Linguisten ja doch [$_{\text{VP}}$Kammermusik spielen]]].
... since linguists PRT PRT chamber.music play.
'...since, indeed, linguists play chamber music.'

These examples show overtly the ambiguity between individual- and generic-readings with respect to the VP as the domain of existential closure. In a similar vein, Borer (2005, p. 138), without reference to German, asserts that "when outside the nuclear scope of existential closure, a variable may be bound by a generic operator, giving rise to a generic interpretation if not otherwise restricted", i.e. when the subject is outside of existential closure as in (65-b). Based on these observations, generic readings are licensed if the subject is in [Spec, IP] and existential closure applies when the subject is in [Spec, VP], binding all variables within the VP.[16]

Relatedly, GEN can be thought of as translating NPs into adverbials that quantify over typical instances or events. This interpretation can avert the issue of variable truth-conditions and subset relationships, as it does not rely on them to be the same across various categories in generic statements. However, it still falls short considering that generics can express unusual properties (66), and that typicality is not sufficient for a generic statement to be true (67):

(66) Mosquitoes carry malaria.

(67) #Books are paperback.

[16] With Diesing (1992, p. 29) attributing this problem to Carlson (1977), note that if the domain of existential closure is the VP, no generic readings should emerge for bare predicates as in *Cellists hate boring bass lines*. I leave the domain of generic predicates open for future research as the focus of this work is on generic subjects.

Striking generics as in (66) are addressed in more detail in Lazaridou-Chatzigoga et al. (2019), but in summary, the findings are that children as young as 4 years old understand that generics are exception-tolerant (although they do not display adult-like behaviour yet). Therefore, these generics are treated similarly to other generics in terms of their behavioural responses, which also show a graded acceptance rate depending on whether they are definitional, majority characteristic or minority characteristics. False generics as in (67), on the other hand, need a separate analysis to explain the lack of acceptance despite the property's prevalence.

As an aside, let us consider some recent research which points towards an under-researched area of the age-of-acquisition component of this theory. Cross-linguistic corpora allow an investigation of the relative order of acquiring generics, adverbial quantifiers (or A-quantifiers, such as *always, generally, usually, often, sometimes*), and quantificational determiners (or D-quantifiers, such as *all, every, most, some, no*). Data from corpus analyses support the view that A-quantifiers are acquired as young as 1;8 years (Kirkpatrick, 2020, for examples in English, German, Italian, and Polish). These data merit a more detailed discussion but for the purpose of discussing their acquisition and their relation to *Generics-as-Default*, it suffices to point out the following parallel: Lewis (1975) argues that A-quantifiers have a tripartite structure, with an operator that unselectively binds any free variables in its restrictor and matrix. This is a close parallel to the proposed modal structure of GEN. In addition to this, Krifka et al. (1995) argue that GEN is closer in form to an A-quantifier than to a D-quantifier. Future research might focus on the gap in the *Generics-as-Default* literature addressing the structure or the acquisition age of A-quantifiers in young children. Kirkpatrick (2020) further points out that based on the observable data points for the acquisition of A-quantifiers and modals (while these emerge as early as 1;8 other related aspects such as TAM features emerge around 2;3), the acquisition of quantification cannot be seen as the barrier in acquiring a proposed GEN operator.

The remainder of this section focuses on indefinite generics and how they can be analysed under an account of GEN. Recall that the proposed structure of a full

DP is $[_{DP}$ D $[_{\#P}$ Num $[_{ClP}$ Cl $[_{NP}$ N$]]]]$. Having introduced GEN, let us reevaluate the structure of generic IS subjects in English. As discussed, these are more limited than BP generics and notably are not accepted in a non-taxonomic reading. Krifka et al. (1995, p. 10, their example (23-d)) exemplify as in (68):

(68) *A lion will become extinct soon.

The semantic denotation of the IS, based on the interpretation given by Borik & Espinal (2012), differs from DNs in (69), where the singular makes reference to properties of objects:

(69) $[\![\text{ tiger}_{\text{singular}}]\!] = \lambda\ x^o\ \exists\ x^k\ [P(x^k)\ \&\ R(x^o,x^k)\ \&\ x^o \in \text{Atom}]$

With respect to an analysis of a IS generic sentence, I propose the following basic structure based on Heim (1982)'s framework (not commenting on the formalism for quantification over possible worlds and situations):

(70) GEN x [x is a tiger $\wedge$ x has stripes]

For a more comprehensive picture, cross-linguistic comparisons should be carried out as they are more informative to understand fine-grained differences regarding the accessible readings. This is parallel to the findings on NUMBER, such that in certain languages, e.g. Yucaltek Maya, classifier and plural marking may co-occur, which is not licensed in other languages, e.g. Armenian. One such example of IS generic subjects could be based on Hindi, which does not have determiners. Dayal (1999, p. 38) asserts that even though a bare singular in object position allows for existential interpretations, "the sentence carries no implication that **only one** book was read" [emphasis added], which can be expected if singular NUMBER referred to one existing book rather than making generic reference to books:

(71) anu kitaab parh rahii hai
Anu book read PROG PR
'Anu is reading the book/a book.'

As a confounding factor, the generic phrase in (71) is in predicate position, which is not the focus of this thesis. These bare objects have independently been noted as problematic by Carlson (1980) and Diesing (1992). By contrast, Borer (2005) mentions that bare nominals, such as *acqua* in Italian, can only have existential interpretation. I propose that this is due to the fact that kind interpretation is determined by definite 'singulars' in languages with determiners, such as English.

In conclusion, GEN can be proposed as an operator that allows for generic reference, but not kind-referring subjects. In English, it can therefore only apply to IS and BP subjects. Thus, GEN can be available if further morphosyntactic layers are present in the subject DP, but not for the minimal NP of kind-referring subjects, which lacks a #P. For the analyses in this thesis, the presence of GEN is not strictly necessary due to the stipulation of distinct DP layers à la Borer and Borik & Espinal. Without stipulating the realisation operator **R**, the differences between generic and individual readings can be explained for by the variation found in English. However, GEN should be borne in mind for cross-linguistic data patterns that were not considered here.

2.5 Connection types and paraphrases

The previous sections have investigated semantic and morphosyntactic properties of subject DPs, both in English and cross-linguistically. They have detailed the subtle differences in acceptability judgements that arise as a result of variation in the underlying forms of the generic (and non-generic) subjects. This section will take a broader perspective and look at the connections that are established between these subjects and the properties that are predicated of them.

To understand the effect of systematic differences in the distribution of subject morphology, we need to develop an understanding of the effect of the connection type between the subject and its property on the subject DP. The main distinction is typically drawn between PCs and SCs. Prasada & Dillingham (2006, 2009) take advantage of this difference and assume two underlying kinds of properties based on a distinction between types and tokens[17]: So-called *k-properties* are roughly analogous to PCs, such as the four-leggedness of a dog while *t-properties* resemble SCs, like the redness of a barn. In other words, "[p]rincipled connections, (i) license the expectation that tokens of the type will generally possess the k-property, (ii) license the explanation of the presence of k-properties in tokens of a type by reference to the type of thing it is, and (iii) license normative expectations concerning the presence of the k-property in tokens of the type." (Prasada & Dillingham, 2006, p. 73).[18]

Referring to the findings from Prasada & Dillingham on the difference between principled and statistical connections, and with a view towards bringing together the hypotheses about differences in connection types raised in this section, the following section will explore the links between language, i.e. the generic subjects that have been presented in this chapter so far, and the concepts that they represent.

2.5.1 Asking the right questions

If we assume that the link between categories and properties is partially determined by the type of category and the type of property, then there must be independent (linguistic) measures to ascertain how they are linked. This applies in a similar fashion to most linguistic knowledge, i.e. it is not immediately obvious to the language user. Nonetheless, a competent user will use it correctly with their tacit knowledge.

Both Prasada & Dillingham, as well as the studies in the subsequent chapters in this thesis, demonstrate that changing the wording of an experimental task can have a significant (and often unpredicted) effect on how the task is interpreted, and therefore the results themselves. *By virtue of* and *one aspect of* are two ways

[17] For a more detailed discussion of the history of the relationship between types and tokens and theories about their connections, see e.g., Murphy & Medin, 1985 on theory-based explanations or Rosch & Mervis, 1975 on prototype-theory.

[18] Note that Prasada & Dillingham make *a priori* assumptions about the type of connection between a kind and its property but the ratings about these connections provided by participants have a wide range, meaning that they may not be the most accurate representations of the targeted connection type. For example, for *Cats drink milk*, the property is assumed to be statistically connected to the kind, yet ratings of the *by virtue of* paraphrases, indicating principled connections, are relatively high. It seems that the properties most strongly representative of statistical connections to natural kinds involve artefacts, while this one refers to diets, which is often associated with principled connections.

of paraphrasing generic statements when such statements are about a kind and a property that is principally connected to it. This extends beyond linguistic analyses of *by virtue of* and *one aspect of* and towards experimental design.

The paraphrases, which are used to target the connection type, can play a significant role in understanding category-property links. Even though both of these paraphrases can be employed to identify PCs, they do so in distinct manners. Empirically, this is noticeable in the consistently different ratings of connection types (as we will see in Chapter 3). Semantically, the differences between these two paraphrases can be broken down when looking at them from a compositional perspective. I will provide a brief outline of some underlying differences between the two most popular paraphrases that are used to target PCs.

The two paraphrases *by virtue of* and *one aspect of* are typically used to determine when a property is principled, as illustrated by Prasada & Dillingham (2006):

(72) a. Airplanes have wings.
b. Airplanes, by virtue of being airplanes, have wings.
c. Having wings is one aspect of being an airplane.

(73) a. Barns are red.
b. #Barns, by virtue of being barns, are red.
c. #Being red is an aspect of being a barn.

These paraphrases often receive uniform treatment, yet they have distinct meanings. This difference has important implications both for psychology, such as Prasada & Dillingham (2006, 2009)'s studies on how these paraphrases are used, as well as for semantics (e.g. Greenberg, 2012) in terms of how we talk about the domain of kinds. I argue that only the former is truly about the construction of the domain of kinds (as opposed to the domain of objects, e.g. Mueller-Reichau, 2013), and that *by virtue of* sets the perspective on how to view a kind. By contrast, *is one aspect of* merely relates kinds and their properties.

Even though seemingly similar, *one aspect of* only speaks to one nominal kind while *by virtue of* relates two nominal kinds. This relation between the two nominal kinds changes what types of properties are permissible. Both *gives live birth* and *has four legs* are acceptable in (74-a), but only properties that are inherited from the superordinate kind are acceptable in (74-b). Further to these constraints, the second nominal, which is part of the *by virtue of* paraphrase, cannot be a subkind of the first (74-c):

(74) a. A dog, by virtue of being a dog, gives live birth/has four legs.
b. A dog, by virtue of being a mammal, gives live birth/#has four legs.
c. #A dog, by virtue of being a poodle, gives live birth/has four legs.

To account for the differences between these structures, one proposition could be that the domain of kinds is such that there are two types of principled properties: those that are principally connected and those that are additionally inherited from a kind further about the kind hierarchy. Inherited PCs are true of a superordinate kind and inherited by their subkinds. Subkinds may inherit properties that are true of these superordinate kinds, such that they are also true of the subkind (e.g. *A dog gives live birth*, though see *#Penguins fly*). In the spirit of Prasada (2016), let us continue to assume that there is a domain of kinds, such that each kind (a) has certain principled properties that are true of it, and (b) is also related to other kinds via subkind relations.

This relates to 'setting the perspective' on how properties are interpreted. Despite their similarities in expressing PCs, and contra Prasada & Dillingham, I argue that *by virtue of* and *one aspect of* do not represent the same kind of relationship between (sub)kinds and properties. They will be analysed in turn.

First, *by virtue of* gives the relevant properties via kind-relationships and therefore picks a higher perspective in the kind-subkind hierarchy. Moreover, a subkind and a kind can be related with a particular predicate *be*. Assuming that generic nominals denote kinds, let us take this *be* (as in *A dog is a mammal*) to have the following denotation:

(75) $[\![be_k]\!] = \lambda y_k \lambda x_k [x_k <_{subkind} y_k]$

where $<_{subkind}$ is taken to relate a kind to itself or any superordinate kind (to wit, *A dog is a dog*, but *#A dog is a poodle*). Under this analysis, both *by virtue of* and *is one aspect of* operate over the property of being a subkind. To account for the differences in (74-b), let us assume that:

(76) $[\![$ by virtue of $]\!] = \lambda P_{<k,t>} \lambda x_k [P(x_k) \& x_k \overset{perspective}{\longrightarrow} P(x_k)]$

Via *by virtue of*, the argument x_k is considered from the perspective of being a subkind of a kind. The output of this function is a kind that is seen precisely from this perspective (i.e. being a subkind of another kind) with respect to the relevant main clause property. In this case, 'setting the perspective' means that a property is considered from the point of view from a superordinate kind. Because of this shift in perspective, the property of being four-legged, which is normally true of a dog, may no longer be true because it is not inherited from the superordinate kind as in (74-b). In choosing the correct perspective, *by virtue of* is a **perspective-setting device** that enables a shift in the kind-hierarchy.

Note also that $[\![$ by virtue of $]\!]$ does not render the properties of the kind because it is an adjunct and not necessary for the grammaticality of the generic expression (77):

(77) A dog (, by virtue of being a mammal,) gives live birth.

The function of *is one aspect of* contrasts starkly with this as it **shifts perspective**. (78) shows that *one aspect of* allows both types of properties (78-a) but unlike *by virtue of* its presence is **not** optional (78-b):

(78) a. Giving live birth/Having four legs is one aspect of being a dog.
b. *Giving live birth/Having four legs is being a dog.

All properties that are true of the kind and the superordinate kind can be expressed because there is no shift in perspective:

$$[\![\text{ one aspect of }]\!] = \lambda P_{<k,t>} \lambda Q_{<k,t>} \forall x_k [P(x_k) \rightarrow Q(x_k)]$$

One aspect of takes a property of (sub)kinds and renders that for all subkinds of a kind the property Q is true of them. *One aspect of* can thus be analysed as a **property-picking device**. In combination with the predicate *be*, it binds the kind and takes a property to apply it to. This yields:

$$[\![\text{ be a dog }]\!] = \lambda x[x_k <_{sk} d_k]$$

Combining these different denotations, its full meaning is the following:

$$[\![\text{ one aspect of being a dog }]\!] = \lambda Q_{<k,t>} \forall x_k[(x_k <_{sk} d_k) \to Q(x_k)]$$

Note that because all subkinds are, by definition, subordinate to their kinds in a kind hierarchy, the individual members of kind belong to the subkind as well. The differences in how the properties of kinds and subkinds are accessed are explainable in precisely these differences between *by virtue of* and *one aspect of*. Reference to subkinds allows the instantiation of relations between the kind and their properties in terms of lattice-structures (Link, 1983).

Including these differences between *by virtue of*, setting the correct perspective for the kind, and *one aspect of*, relating kind-properties, allows us to refine the notion of PC generics. This is an extension of Greenberg (2012)'s analysis of "(non-)accidental" generics, in particular her analysis of *happens to*-generics, which is then compared to *by virtue of* generics. One difference of Greenberg's formalism (following Brennan, 1993) that I do not address here is that she assumes possible worlds with normal individuals that may have a *salient property*, S_c, as opposed to any arbitrary property P.

In addition to that, it also follows the assumption that the presence of NUMBER enables access to subkinds or individuals, while at the same time blocking direct kind access, as introduced in Section 2.3.3 above. This is in line with other kind-oriented approaches to genericity, the analysis outlined here can further extend to previously observed morphosyntactic differences between kinds and subkinds (Borik & Espinal, 2012, 2015). The next section will elaborate on the approaches of category-property links and their acceptable paraphrases by taking a step back and considering the nature of categories per se and the role of linguistic input in their formation.

2.6 The link between categories and language

To develop an understanding of the interface between categories and language, I will discuss some general theories of category denotation. In its vaguest meaning, let us think of a category as a grouping of features. Following Murphy & Medin (1985, p. 289), who use the terms 'concept' and 'category' mostly interchangeably, a category is formed of a group of features that are "informative, useful, and efficient" as opposed to "vague, absurd, or useless". These features can be perceptual or causal (thereby providing a theoretical understanding), but ultimately, the precise nature of these features is as of yet inconclusive. Based on these definitions, Carey (2009, Ch. 13) comes to the conclusion that concepts are "mental symbols." Before looking at language-category links in more detail, we may take stock of what has been relatively firmly established: perceptual similarity, correlations of attributes, and causal theory each are not sufficient in their explanatory power (Murphy & Medin, 1985 referring further to Rosch, 1978; Smith & Medin, 1981; Mervis & Rosch, 1981, among others). To foreshadow some of the upcoming discussion, consider that not having enough explanatory power might simply indicate that these theories selectively apply under certain circumstances or only for specific types of categories. In other words, it is possible that no single theory is or must be able to account for all category meanings and categorisation processes.

The question arises what it is specifically that a concept denotes. One popular theory of concepts is that they are denotations of their lexical items (see Murphy, 2002, 2005). This implies that the meaning of a concept is directly expressed through its lexical meaning (although see Carston, 2012 for an alternative account). Löbner (2002) and Barsalou (2000) argue that concepts are not equal to kinds, but that kinds are more abstract and reliant on the descriptive content of concepts (as presented in more detail in McNally, 2017). McNally (2017, p. 41), however, comes to the conclusion that "[k]inds ... serve as a bridge to connect referential and cognitive or conceptual approaches to meaning." While I agree that concepts, categories, and kinds are all closely connected, I assume that nouns denote properties of kinds and that they therefore express conceptual meaning, following the discussion in

Section 2.1. Whether this entails that abstract kinds are the linking point between referential and cognitive approaches to meaning, or whether this is something better carried out by concepts, I shall leave to other researchers. I acknowledge though that a later discussion on whether concepts are represented as prototypes or mini-theories (see Chapter 6) complicates the picture, especially when stipulating that multiple theories of conceptual structures could be combined.

Thus, while these are important considerations, I will return to discussing them in more detail in Chapter 6 within the context of the experimental findings of this thesis. At this point, it is more pertinent to consider the main interaction that has sparked the interest in this line of research. Based on the above question of how we acquire knowledge about something that only exists in the abstract, we must acknowledge that the brief (and partial) answer is that this type of knowledge is passed on through language. This is underpinned by the fact that no child or adult can investigate the state of the world through personal experience alone. Relatedly, we have conceptual knowledge that extends beyond the current state of the world, and we are unable to check past and future states. Yet, knowledge about counterfactuals (*If Fluse were a chicken, she...*) and past states (*The dodo is/was a flightless bird. But yes, the dodo is extinct.*) influences our generalisations. Moreover, I have only referred to kinds that currently have or previously had instantiations in the world here (my dog Fluse, the bird whose skeleton cast is exhibited in the Natural History Museum in Oxford). For abstract kinds, such as LOVE or TRIANGLES (or LOVE TRIANGLES), there are no instantiations at all.

The fact that we are still able to talk about these concepts and members thereof underlines that without linguistic input about what the world is or could be like, we could not make any of the generic statements we confidently and frequently make. This further supports tight links between linguistic and conceptual representations. Sections 2.1-2.3 argued that (covert) morphosyntactic variations determine whether reference is established directly to a kind or whether reference is made to an arbitrary individual. I extend this general idea to a more abstract level but I do not argue that category formation is exclusively based on linguistic information (unlike direct kind reference vs generic reference to a group, which is strictly reliant on the linguistic information conveyed in the subject DP). Instead, establishing concepts relies on a range of information, which can be explained by drawing on information provided by linguistic labels (these phenomena include, e.g. labelling effects). Based on the limitations in the types of properties certain generic subjects can express, it seems too simplistic to assume that there is no, or only a limited, interaction between concepts and linguistic cues.

Carlson (2010, p. 16) aptly summarises this intricate relationship between language and concepts as a "striking confluence of interests between the study of concepts in the psychological and cognitive science literature, and the study of certain types of sentences in the formal semantics literature." This thesis equally focuses on linking up the advances made in the psychology literature with those in the linguistics literature. Carlson initially sees these confluences as a hindrance to understanding the pieces that make up the generic puzzle. This is an understandable concern as one of the notorious problems when talking about category-property links is that the term 'concept' itself is still under debate.[19] Evidently, without an understanding of the nature of concepts, it is difficult to describe the relationship that they have with language, as indirectly observed by their linguistic distributions. For further investigation, I maintain that these two components mutually influence each other and can be used to gain insight into their respective natures. Similarly, Carlson (2010, p. 21) also concludes that "the hazy confluences (...) between talk of concepts and generic sentences is not coincidental, and that the nature of concepts can play a serious role in determining how some generic sentences are understood." Drawing on both linguistic and psychological literature (Carlson, 1977; Greenberg, 2003; Cohen, 2005; Prasada & Dillingham, 2006, 2009), Carlson illustrates how various proposed mechanisms ultimately seem to target the same underlying structures.

[19] I will follow Carlson (2010) in using the 'subjective', as opposed to 'objective' interpretation of the term *concept*. This is the usage normally assumed in psychology, which contrasts with the more abstract notion adopted in philosophy. Under the 'subjective' interpretation, Carlson continues, concepts are seen similarly to words, where it is close to impossible to assume that two individuals share the exactly same interpretation (or lexicon) but are able to converse and use them mostly without problems.

Recall the notion of *k-properties* and *t-properties*, introduced above in Prasada & Dillingham's work, as Carlson relates it to other observations, i.e. reframing 'essential', 'stable' and 'suitable' properties as *k-properties*. One early example in Carlson (1977) is particularly relevant to this discussion. It not only illustrates the difference between these two properties but also the effect of category knowledge, in this case artefacts as opposed to the often-cited animal kinds. Carlson (1977) uses minimal pairs to show that (79-a) is acceptable, but that there exists an asymmetry between the two closely related (79-b) and (79-c).

(79) a. Cars have tires.
b. #Cars have black tires.
c. Cars have air-filled tires.

These acceptability judgements cannot simply be based on subset-relationships or prevalence of the property. It turns out, as Carlson argues, that the difference is that being black is a *t-property* of tires, defined as irrelevant in terms of function for an artefact kind, and being air-filled is a *k-property* of tires, a necessary prerequisite for function of the artefact car.

A similar observation has been made about different acceptability judgements depending on whether the kind is well-established or not (presented by Carlson, 1977, attributed to Partee, 1974):

(80) a. The coke bottle has a long neck.
b. #The green bottle has a long neck.

This example illustrates how different types of kinds allow for different types of properties. These differences in acceptability **within** generic subjects are highly important as this thesis is mainly concerned with the interaction of morphosyntactic cues and properties. While (80) shows that DS subjects, too, are sensitive to acceptability limitations based on the kind itself, other examples have shown that these are interactions between the form of the subject and the property type. While BPs can express both *t-properties*, or properties that are only statistically related to the kind, as illustrated in (81-a) and (82-a), IS subjects are preferentially used to express *k-properties* (see (81-b), contrasting with the unacceptable (82-b)):

(81) a. Madrigals are polyphonic.
b. A madrigal is polyphonic.

(82) a. Madrigals are popular.
b. #A madrigal is popular.

Supported by these minimal pairs, I follow previous researchers in the assumption that both the content of the subject, representing a concept of some sort, and the content of the property itself, restrict how the two can be combined.

In terms of terminology, I will refer to *k-properties* as principled properties (PCs) and *t-properties* as statistical properties (SCs). However, I am conscious of related discussions about the nature of PCs, and, for instance, whether they require essentialism. While *k-properties*, as proposed by Prasada & Dillingham and further discussed by Carlson (2010), encompasses a broader notion of properties, I prefer to use a more narrow definition to be able to tease apart measures of essentialism and connection types. This is important for some of the experiments presented in subsequent chapters. This is not to say that ultimately, PCs are distinct from *k-properties*, or that *k-properties* might not be a better denominator for the theoretical and empirical data presented here.

Prasada & Dillingham's studies stand out because they do not focus on the linguistic aspects of generic language per se, but are interested in the underlying conceptual distinctions they denote. For their research question, language is merely one way to access these abstract representations (2006, p. 78). They use to their advantage the fact that English BP generics can express both principled and statistical connections. In their 2006 study, participants were asked to provide both a prevalence estimate (in percent) and to judge a paraphrase (on a Likert Scale) of a previously presented BP generic. This was done with paraphrases as introduced above in example (74), repeated here as (83):

(83) a. A dog, by virtue of being a dog, gives live birth/has four legs.
b. A dog, by virtue of being a mammal, gives live birth/#has four legs.
c. #A dog, by virtue of being a poodle, gives live birth/has four legs.

An important asymmetry lies in the observation that while SCs do not necessarily presuppose an underlying PC, PCs are typically also statistically prevalent.[20] Both of these connection types can be expressed by BP subjects, which consequently leave the connection type underspecified. However, as alluded to above, the category type itself can guide our expectations as to whether a connection is principled or statistical. Prasada & Dillingham (2009) provide evidence that they differ conceptually: PCs were more accepted when participants were asked if the object possesses the predicated property, by virtue of being the kind of thing it is. Prasada & Dillingham's findings cumulate in their 'Aspect Hypothesis', which states that "to represent a principled connection between a kind and a property requires representing the property as one aspect of being that kind of thing" (2009, p. 407). Differences in licensing of *k-properties* can be explained by differences in the kind of connection an object has with its property, where PCs license normative expectations and support formal explanations (p. 429).

Let us return to the main question of this section about the nature of concepts and the role of language in understanding their formation and structure. Even when we maintain that an interaction of linguistic cues of the category and the type of property is essential to our understanding of concepts, clarification on their nature will ultimately be essential for understanding their interaction. One such context in which this becomes more obvious is that their meaning must be able to account for complex and ad-hoc concepts as well. So far, the examples in the literature have

[20]Though consider cases such as *Mosquitoes carry malaria*, where less than 1% of them do. These are so-called 'striking generics' (e.g. Khemlani et al., 2012; Leslie, 2008) that are judged true despite their low prevalence. The type of connection here might be of yet another type, causal, which I will not discuss here (but see Prasada & Dillingham, 2006, among others). Others have also suggested thought experiments (e.g. Mueller-Reichau, 2013) in which an accident affects all instances of a kind, such that a principled property is no longer statistically prevalent. This might be, for example, an accident rendering all dogs three-legged: *All dogs are three-legged* is true, yet the truth-value of *Dogs are four-legged* is not affected.

focused heavily on simple, mostly monomorphemic categories. This could, in the first instance, be a reflection on their conceptual accessibility or frequency.

Yet, any theory of concepts must then either be compositional or take seemingly complex concepts to have an underlyingly simple form. The latter seems undesirable in the face of subkinds and modified kinds, especially given the analysis of subkinds in Section 2.3.3. Moreover, Carlson (2010) points out the concern that monomorphemic names would already be rendered unlikely due to concepts such as *the Indian elephant*. If we assume that nouns are names for (properties of) kinds, and kinds are the basis for concepts, then we could return to the covert morphosyntactic layers introduced in Section 2.3.1 above. Under this account, the different interpretations between *elephants* and *the elephant* are not due to the compositionality of multiple words, as the even the morphology of *elephants* is in itself complex. Recall that the difference between a generic and an episodic interpretation of *(the) elephant* is based on whether NUMBER is present as [singular] or absent entirely.

Consequently, the former theory is a more likely candidate for the make-up of concepts. Under this view, kinds and concepts are closely aligned and further, this aligns with the theory that in its most basic form, nouns refer to kinds and any object-level or individual-level reference is achieved by additional layers in the syntax. In a parallel fashion, basic categories can be taken as the building blocks of the conceptual system. In this case, phenomena such as labelling effects Xu (2002); Jaswal & Markman (2002); Waxman (2003); Gelman (2004); Graham et al. (2004) can be accounted for in a straight-forward manner by relying on the underlying kind reference evoked by the usage of a category label.

Some theories aim not to conflate the meaning of nouns as kinds with the meaning of concepts or categories. I aim to make their main differences explicit, although this issue is still under debate and a fully distinct definition of what kinds are, as opposed to concepts, has not yet been agreed upon. For the current purpose, kinds remain abstract entities that are independent of any (possible) instantiations. It then follows that the meaning of a kind is dependent on the lexical information, expressed by the semantic representation, of that noun. This has been independently raised by researchers arguing that concepts of emotions do not form a (natural) kind (first discussed in Griffiths, 1997, but see Griffiths, 2002 for a more recent account). One example that illustrates this point is when scientific understanding illuminates the underlying nature of a kind, which might not affect any non-specialist's concept thereof. Griffiths (2002, p. 2) points out that "[v]itamins are not, as was once thought, 'vital amines' but a diverse group of chemicals (...). So the concept of a vitamin can be analysed, and individual vitamins and even some groups of vitamins are natural kinds, but 'vitamin' itself is a superficial descriptive category." Surely, a scientist's categorisation processes and membership criteria are affected by a better understanding of what a vitamin is, yet the general public's conception, and probably the scientist's everyday usage of the term, remain largely the same. It also means, by extrapolation, that a kind can be dependent on the information expressed by a category.

This idea that concepts are not kinds is also picked up in Machery (2009) who argues further that the kind 'concept' is multi-faceted, or not homogenous. This clearly delineates kinds from concepts as it evokes the kind 'concept', and vice versa allows for a conception of 'kind' that may vary between individuals (for a broader discussion on natural kinds, see Bird & Tobin, 2018). The fact that this distinction is possible sets these two terms apart.

Then, concepts themselves can then be seen rather as mental representations of a kind, which are evoked by the meaning of the noun that expresses the corresponding kind. These representations could be interpreted as *mental images* or *ideas*, as proposed by e.g. Locke (1690) and Hume (1739). This view has been modified in favour of more structured theories, that have developed into *Language of Thought* (see Fodor, 1975). In this and subsequent work (1981; 1998), Fodor explains the now well-known 'Mad Dog Nativism', which, in its strongest version, claims that all concepts are innate. Rey (2014, p. 110) further defends this viewpoint from a less strict sense by arguing that at its core, the 'Mad Dog View' provides "conceptual Atomism", which might be innate and learned. This is both promising (in that it combines the advantages of various conceptual acquisition theories) and problematic (in that it needs to address the shortcomings of the theories as well as how they interface; this is analogous to proposing both UG/nativism and statistical learning for first language acquisition).

More closely aligned with *Generics-as-Default* are theories about concepts in the cognitive sciences in which concepts are seen as the default mental representation, as proposed by Pinker (1994, 2007) and subsequently by Fodor (e.g. 2003). As the ontology of concepts is still negotiated, let us assume that minimally, without mental concepts, categorisation processes are impossible. Focusing on the structure of concepts, as opposed to the ontology of concepts, I will discuss the the differences between **prototype theory**, **exemplar theory**, and **theory theory** in more detail in Chapter 6, with their respective points of relevance to the studies presented in experimental chapters.

To finish this section, recall that I began the investigations in this thesis by illustrating that kind reference, as provided by numberless definite DPs, is special. I have shown that these subjects combine with kind-level predicates and are sometimes preferred over other generic subjects. These other generic subjects make generic generalisations which are about categories in the world, and therefore build concepts in the human mind. Seeing that in generic statements, these are often used interchangeably, this distinction can be borne in mind for further research on this particular distinction. As my research relies on the distribution of linguistic form and the expectations as to the connection with the property, I will use the terms 'category' and 'concept' interchangeably, unless a finer distinction is necessary. The term 'kind', however, will be used specifically to distinguish direct kind reference, as given by the DN subject of a generic statement.

Having provided a broad overview of how concepts and language influence one another, the next two sections will first look at category induction theories under the lens of the two main types of categorisation strategies discussed in the literature (Section 2.6.1). Section 2.6.2 will develop the argument that assumptions about category structure affect the linguistic forms available to express the generic subject and vice versa.

2.6.1 Theory-theories and prototype/exemplar-theories

The fact that the human conceptual system allows abstraction from instances to kinds in the first place is crucial for novel category induction but, likewise, we extensively employ knowledge about existing categories to assist with the induction of novel ones. One of the main questions about induction concerns the types of features that specific instances possess, which are abstracted away from to form categories. The 1970s had a surge of research on prototype and exemplar theory (e.g. Rosch & Mervis, 1975; Rosch, 1978; Mervis & Rosch, 1981), which lasted into the 1990s (e.g. Wisniewski, 1995). Experimental studies that were carried out to gather empirical support for these proposed theories targeted the mechanisms that underlie category-based inductions. One way of doing so is by presenting participants with an instance of a category (often labelled the 'target') and asking whether another instance was also a member of the same category.[21] These are known as 'match-to-sample' studies, which I will discuss in more detail in Section 2.7 in light of the novel data presented in Chapter 6.

Under the lens of prototype theories, the main assumption of match-to-sample studies is that exposure to an instance of a kind will reinforce expectations about their crucial properties. Seeing that concepts often have 'fuzzy boundaries' and categories such as GAME are only coherent if a family resemblance is assumed, their appeal becomes obvious (see Wittgenstein, 1953, for example). It could also be argued that there is a cline within frequentist theories, such that a concept is initially formed based on exemplars and that, over time, with more exposure, these shift into prototypes, granted that more exemplars are accessible to the conceptual system. Geeraerts (2008, p. 146f) lists four main properties of prototype theories (paraphrased here):

1. Categories under a prototype theory do not rely on a set list of criterial properties

[21] For a more detailed overview of the role of induction tasks and existing knowledge, see Murphy, 2002, specifically Chapter 6.

2. Categories under a prototype theory have a structure of family resemblance
3. Categories under a prototype theory have various degrees of category membership
4. Categories under a prototype theory have blurred, or 'fuzzy', boundaries

Exposure-based theories contrast starkly with theory theories (with the first comprehensive overview provided by Gopnik & Meltzoff, 1997) in that a sort of inherent essentialism of the kind is assumed for the latter, thereby categorising based on underlying mini-theories (see e.g. Carey, 1985, 2009; Keil, 1989). Carey (1988) in particular provides an example of a child's concept of *cats* vs *baby cats* which is not analogous to *worms* vs *baby worms*, but instead the contrast between *worms* and *short worms*. Exposure-based theories could not account for the fact that a child distinguishes between different types of smaller and younger animal offspring other than proposing an underlying difference in their essence, which, moreover, changes once children become adults, seeing that the types of exposure do not change throughout their lifespan. Theory theories, presupposing a definitional content, are also particularly well-suited to defend examples of transformation stories in which people maintain that an individual is still the same even if superficial characteristics have changed. On the other hand, being able to revise an understanding of a concept, with the input remaining constant, is in line with revising a theory by restructuring and reframing how the data points are understood. In a similar vein, Prasada (2016b) has argued that transformations are easily accepted if an individual changes into another instance on the same level of the kind hierarchy (e.g. from a frog to a dog), but not another instance of the same kind (e.g. from one dog into another dog). Equally, one could image that these transformations are less accepted if they occur on different levels of the kind hierarchy (e.g. from a blue whale to a pygmy blue whale). Theory theory, being particularly apt at handling conceptual change within a life span is also a preferred theory to explain conceptual change of an instance as perceived by another. While not the topic of this thesis, they also seem to be more appropriate than exposure-based theories to explain conceptual change over longer stretches of time as concepts under this theory are not determined by external instances but their non-obvious, underlying characteristics.

As Osherson & Smith (1981) and Landau (1982) have suggested, prototype theories might be combined with theory theories to provide a full account for the structure of concepts. This view is not without problems, as hybrid theories are often assumed to shift the problems of either of their shortcomings instead of solving them. For instance, what does the theory-theory component of a concept encompass in terms of non-empirical input? From the perspective of exposure-based theories, the question of how outliers and irrelevant data input are dealt with remains. I hope to show that different linguistic cues may guide the conceptual towards one conceptual route over another. In assuming that various theories are involved in category formation, in addition to the linguistic cues, the preferred conceptual theory might also be dependent on the type of category at hand.

Let us briefly entertain a toy idea to illustrate what this dual-route might look like: We could assume that artefact categories are more reliant on prototypes, due to their teleological nature. They are more often than not purpose-built and follow an intended structure. Their form is often determined by carrying out its purpose. By contrast, animals and mathematical objects could be assumed to be more dependent on theory theory based on their high essentialism and definitional nature. Theory theory does not presuppose that a fully-fledged account of the category's inner workings is provided. Instead, as Murphy & Medin (1985, p. 290) put it, "any of a host of mental 'explanations"' suffices to count for a theory theory of concepts.[22]

In addition to differences in the conceptual ontology based on the nature of the category and the linguistic cues with which it is evoked, one could also draw upon personal differences. One such idea is the "personal construct account", which argues that statistically induced concepts, i.e. prototype theories (and, to

[22] This is incidentally in line with popular science findings on the illusion of explanatory depth (see e.g. Rozenblit & Keil, 2002; Keil, 2003). This effect describes the phenomenon that we are under the impression that we know how things, especially everyday items work, but when put to the test, we have very little knowledge about how things actually work. Instead, we only have what Keil calls "skeletal frameworks" that offer rather shallow explanations as opposed to the detailed knowledge we expect to have.

an extent, exemplar theories), are not accessible to the same level in people with schizophrenia or bipolar disorder. Limiting the effect of information due to their statistical prevalence then limits the set of beliefs that is available to a person. Stam (1998, p. 187)'s research addresses the idea that "[l]inguistic and social communities are essentially the repositories and generators of knowledge even as we are individual knowers." Stam lays out, following Kelly (1955) and Mancuso (1996), that if most of the linguistic, and perhaps visual, input appears to be random, no abstract patterns can be detected, thus limiting the construction of categories and concepts. This is the *problem of coherence*, i.e. "how can the linguistic inputs ever account for the development of constructs unless the constructs already contain something of what is in language?" (1998, p. 193). However, as I investigate the role of morphosyntactic variation in healthy adults and children, I will leave this consideration for future research.

Based on these initial observations, whether we assume that concepts are built based on knowledge-driven or exposure-driven theories, it seems that there are limits as to the types of categories we naturally assume are possible. The properties around which they are centered may be non-obvious or complex, more fine- or coarse-grained, but without additional input, there are certain biases, such as a shape, taxonomic, or whole-object bias (see Markman, 1994 for further detail on these three assumed biases from a developmental psychology perspective), that guides the conceptual system and explains a large variety of the categories we stipulate.

I have intermittently referenced the significance that pre-existing (superordinate) category knowledge can play when constructing novel concepts. The next section will address this more explicitly by looking into the assumptions we might have for the same type of property if predicated of different types of categories.

2.6.2 Checking your assumptions

I will start with the observation that the same property can be interpreted in different ways if it is predicated of one type of category, say an animal kind, as opposed to another, say an artefact. This section will lay out our underlying expectations of how a property might be connected to a category based on our knowledge of the world. An understanding of the underlying semantics of the generic subject, based solely on the syntax-semantics interface, is not sufficient on its own. We have already seen that an understanding of the category itself is not sufficient either. Then only the interaction between these two can ultimately account for the way in which we quickly acquire world knowledge as children and novel category knowledge throughout our life.

Earlier, I introduced theories of noun denotations and category ontology that support the idea that this interaction is necessary. Before moving on to the a developmental perspective, it is worth repeating one particularly poignant example. Returning to *t-properties* and *k-properties*, in Prasada & Dillingham's 2006 studies, *being red* is used as a principled property, or *k-property*, for cherries, but as a statistical property, or *t-property*, for barns. This shows that only through a combination of linguistic and conceptual knowledge can these be distinguished as different types of properties, which in turn allow for different subject types. This is evidenced by acceptability judgements of various paraphrases that target PCs and SCs. We must therefore assume that the predicate itself is not categorised as one specific type of property.

This will become relevant for the upcoming learning studies which train participants on novel kinds with novel properties. They differ in the extent to which participants know what type of kind they are learning about, i.e. whether they are shown a picture of an animal or given no linguistic and non-linguistic information about the type of category. The results must be interpreted keeping in mind the underlying assumptions participants already have, as they will have developed an understanding already as to which types of properties are normally principally or merely statistically connected to an animal kind.

Relatedly, consider the role of assumptions about property types with regard to experimental design. When talking about PCs, I assume that properties such as fur type and colour, number of extremities or reproductive means are representative examples of properties that are principally connected to animal kinds. While this is true on first impression, this overlooks the fact that there might be equally obvious but less explicitly discussed properties. The one that I will focus on is the role of shape. In some of the experiments, the stimuli are designed by pitting a *shape* against a *property* that is explicitly predicated of the novel animal kind. This *property* is still either principally or statistically connected to the animal. However, by assuming that we can manipulate a principled property against the animal's *shape*, we are led to two options. Either, *shape* is inherently different than a principally connected property, and by extension not a property in its own right. Or it implies that shape is **also** a principled property, in which case we pit two principled properties against each other, one of which is explicitly mentioned and the other one is implicitly represented.

I have overtly focused on some assumptions that are often not teased apart explicitly when investigating how categories are learned. The main take-home message of this section is that it is difficult to disentangle these confounding factors. A revision of the implicit assumptions is crucial to target the individual variables involved and their significance in any learning studies. While they will not all be separated out in the experimental designs below, I will focus on the contribution that any pre-existing knowledge could have played in the development of category-property links for the novel kinds. The next section will provide an overview of the developmental aspect of categorisation, focusing specifically on children's ability to generalise based on limited personal experience with the world.

2.7 Development

Developing categorical knowledge about the world is a joint venture between word learning, non-linguistic gestural input, and the emerging understanding that the world is organised in categories and not simply a continuous stream of disjointed stimuli.

I presuppose that basic linguistic competence is present in interlocutors of the examples discussed and, later on, in the participants of the experiments. With regard to the development of this competence, the significance of gestures and joint shared attention is discussed in e.g. Woodward (2003) and Tomasello (2003). Further research in language acquisition has demonstrated that first words are typically tangible objects and people (as well as subject to phonological constraints), which are commonly found in a child's environment and thus familiar to the child (discussed in more detail in Saxton, 2010). Once these are acquired, the interplay and mutual influence of category formation and linguistic representations increases. This subsequently enables children to recognise individuals as category members, and crucially allows for a bidirectional understanding. In other words, children learn that there exist individual differences within category membership (e.g. that two individual German shepherds are both members of the category German shepherd) and that there is differentiation of two (near-)identical instances of a category as two distinct individuals (e.g. two paper clips from the same box).

This developmental feat continues past the initial acquisition period and is a guiding mechanism for category acquisition. There are five general cognitive biases that are instrumental: (1) the whole object bias, (2) the mutual exclusivity bias, (3) the taxonomic bias (all of which are laid out in more detail in Markman, 1994), (4) shape bias, and (5) basic level bias.

The whole object bias picks up the Quinean conundrum, as famously introduced in Quine (1960)'s example on the "indeterminacy of reference." Without any further instructions, a new label for an object is assumed to refer to the whole object, not just parts of it, e.g. the novel word *gavagai*, as uttered when pointing to a rabbit, is taken to refer to the whole rabbit, not just its ears. This bias is helpful in understand developmental stages in category and word learning, but will also be assumed for the picture book studies in this thesis.

The second proposed bias that guides the development of category knowledge is the mutual exclusivity bias. Simply put, it states that there is only one label for each object (Markman, 1994). Thus, any new label should generally be attributed to a new object. If a child has acquired the label *hat* to refer to hats, then upon hearing the novel word *vase*, the child will not assume it is an alternative label for hats, but pose that it is a label for another object, in this case a vase. Vice versa, after learning the label for hats, upon seeing a vase, a child will not assume that *hat* also applies to *vase*. I caveat this by pointing out that young children typically over- and underextend labels during first language (L_1) acquisition. This means that they often use one label, which they will have acquired already, incorrectly to refer other novel objects. I will not take a stand as to whether this is due to the fact that they assume these are instances of the same category, or whether this is due to not having the correct means of referring to a novel category and thereby resort to a label they already acquired. As I mentioned above, this thesis is concerned with the role of linguistic input in category acquisition once L_1 is (mostly) acquired.

The taxonomic bias describes the underlying assumption that a label, in a context of ambiguity displayed by instances, is interpreted as referring to members of the same (superordinate) kind. The taxonomic bias thereby highlights category-specific knowledge, and places less emphasis on alternative means through which objects can be related, such as thematic relationships or similar shapes. Imai et al. (1994) pitted these three types of potential biases that are hypothesised to influence categorisation in a series of experiments. A popular method used in studies targeting categorisation, ultimately designed to understand how world knowledge is acquired, is a match-to-sample study (mentioned in Section 2.6.1). Participants were shown a target image and were asked to decide on a match between three further images (see Figure 2.6). Overall, participants matched taxonomic matches more often than either of the other two options. This means that for this trial, participants were more likely to choose a pie when the target image was a birthday cake, than they were to choose a hat (similar shape) or a present (thematic match). Note that this experiment introduces objects from various taxonomies, while other match-to-sample studies often stay within on type of kind.

Fourth, seeing that a taxonomic bias is unnecessary once categorisation takes place within one taxonomy, other biases can become more important. The *shape* bias in particular aids grouping objects together if they are similarly shaped (as was the hat compared to the birthday cake), which is a more powerful categorisation mechanism than colour, size, pattern, material or texture (as studied by Landau

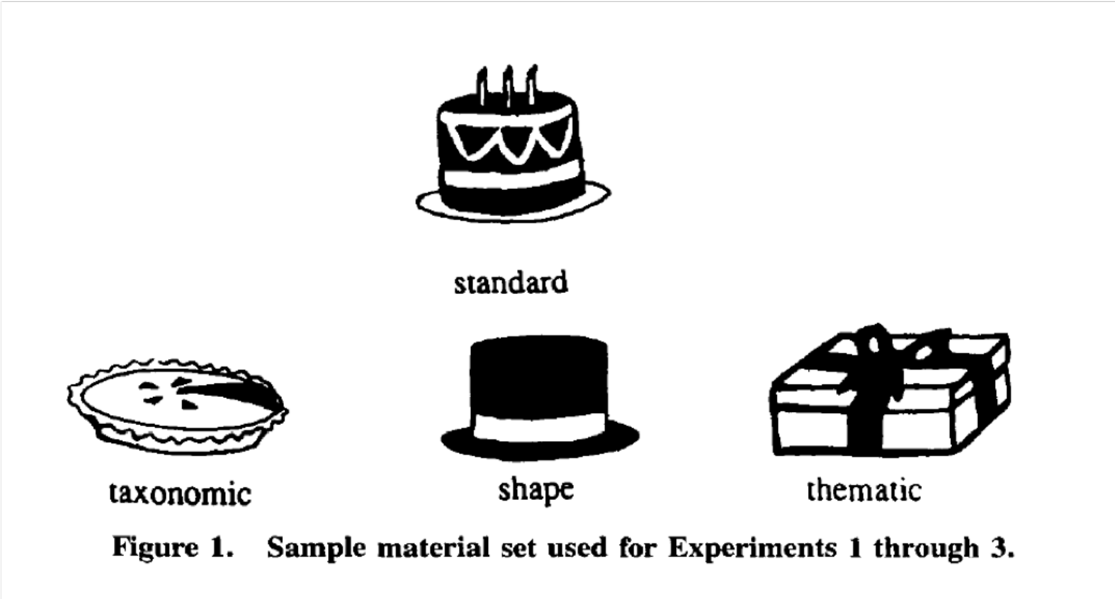


Figure 2.6: Sample material of Imai et al. (1994)'s match-to-sample study, pitting taxonomic, shape, and thematic matches.

et al., 1988). Therefore, studies about novel animal kinds might assume a shape bias in their stimuli design.

Finally, the basic level bias illustrates that categorisation initially targets basic level concepts. For instance, the concept BIRD is easier for children to remember, and by extension to memorise, than the subkind ROBIN or the superordinate kind (or class) AVES. One piece of supporting evidence is given by the make-up of a child's early vocabulary. Developmental and cognitive psychology further stipulate that the underlying reason for the prominence of basic level kinds is their level of information density, which is further related to the level of abstraction or specificity that is necessary to distinguish and group instances.

These general biases guide categorisation before and during concept acquisition, as well as after concepts have been established and novel objects are being evaluated as to their category membership (e.g. atypical designer chairs, which still are members of the category CHAIR).

Further, these biases, which persist in category acquisition, are important for understanding how novel objects are given category membership in the conceptual system. However, if we focus on the role of linguistic input, the influence of these biases needs to be contained as much as possible. For the remainder of this section, I will lay out how the influence of these biases can be reduced such that the effect of variation within generic subjects, and the effect of their different semantics, can be maximised. I base this on reprises on the observations above that linguistic development and categorisation develop in lockstep, and advances in one area aid the understanding in the other area. This leads to an entangled understanding of concept acquisition and helps in establishing what the contributions of different linguistic, or more specifically morphosyntactic, cues are.

Relatedly, this explains how this topic has interested linguists, philosophers, psychologists, and cognitive scientists. Despite their joint interest in kind-referring and generic statements, a major challenge of this research now lies in reconciling the various contributions stemming from the vast and disparate literature. Psychology studies typically approach genericity by contrasting generic with quantified or specific statements to address the special status of generic language and knowledge in child development. By contrast, linguistic theory has primarily been concerned with accounting for the variable truth-conditions of generics and has proposed modal and quantificational analyses.

The main puzzle that the linguistic aspect of generic statements raises for scientists is how we acquire generic expressions initially, i.e. its developmental aspect. Unlike other types of linguistic expressions, such as quantified or specific statements, they do not have an overt operator that distinguishes generic from individual reference. This relates back to the initial observation that our personal experience is limited such that we should not be able to make statements about kinds as a whole with confidence, e.g. assert that *Tigers have stripes* without seeing many instantiations of tigers. A more poignant example would be *Oxygen is the third-most abundant element in the universe*, given that oxygen is also invisible or *The dodo is a flightless bird*, given that dodos are also extinct.

An analysis of the hypotheses and methodology of previous experiments addressing this “generic puzzle” unveils important variables that can be better manipulated from a linguistic standpoint. The aim is to arrive at more representative results that can account for the sometimes-conflicting results from linguistics and psychology by taking seriously the three different types of generic subjects in English, indicated by their distinct linguistic make-up.

Using existing categories, be it for acceptability judgements or as part of the experimental design, will always presuppose some sort of category knowledge. Even stipulating novel categories with accompanying pictures evokes preexisting superordinate kind knowledge. One more radical way of stripping generic statements of category knowledge is by only maintaining the structure of a generic statement, but replacing all content words with nonce-words. With such a design, the effects of using a plural marker, a definite or an indefinite determiner can be compared directly and not as part of conflated assumptions about what certain categories and their properties are like. I will explore this in Chapter 3 below. Note though that this focuses less on linguistic development during language acquisition and more on applying the tacit final-state knowledge to novel situations that imitate language acquisition.

In summary, I have argued above that generic interpretations are built compositionally from the subject's morphosyntactic pieces. Their distinct semantic interpretations are expected to lead to conceptual differences, observable in distinct behavior. Even in the presence of other guiding principles, children must come to distinct representations for the individual morphosyntactic pieces of English generic subjects. Ultimately, these representations should aid the conceptual system. In other words, by better understanding how morphosyntactic knowledge can be targeted and learned, we also better understand how complex linguistic representations lead our conceptual system to various distinct generic interpretations.

As the final piece in this overview, let us move from general assumptions about development to the input that a developing child is exposed to. Section 2.8 presents a corpus analysis of the CHILDES database, looking into the morphosyntactic variation in generic subjects that children are exposed to. Only by knowing what type of generic language children are exposed (and by estimating their proportions) to can we ultimately determine whether frequentist accounts are an adequate theory for the acquisition of generic subjects.

2.8 Corpus analysis

The puzzle of how children acquire a generic interpretation for BP, IS, and DS subjects that reflects both their joint status as generic subjects and also their distinct semantic distributions in the absence of an overt generic marker remains to be solved. Here, I will complement general developmental considerations raised in Section 2.7 by analysing child-directed speech as it occurs in every day interactions. This should provide us with a better idea of the linguistic structures that children hear as well as the contexts in which they occur. This knowledge will inform the interpretation of experimental findings of child participants' data. My main motivation is to use data from naturalistic settings in corpora and complement these findings with experimental work that targets specific aspects of language acquisition. This section will provide a more holistic answer to the question of how we come to understand when generic reference is intended.

Bearing in mind that the general aim of corpus studies is to find a representative sample of the language and the associated linguistic structure in question, a corpus analysis yielding representative results must include corpora that can account for the variety and variability of a speaker community (Biber, 1993).[23] I anticipate that the variation of generic subjects types will be low. Native speaker intuitions and the literature, specifically experimentally stimuli, strongly favour BPs. As such, I hypothesise that this subject will also be highly prominent in the corpus data. DS subjects, which express kind reference in English, are expected to be particularly sparse. Previous research by Gelman et al. (1998) has demonstrated an overall high prevalence of generic statements in child-directed speech, at least in parent-child interactions. However, this observation is based on looking at BP subjects as generic subjects, and does not consider singular generic subjects.

[23] This is complicated by the fact that an analysis that is comprised of multiple corpora normally includes corpora that vary in sample size and transcription accuracy. I thank André Eliatamby (p.c.) for his valuable comments on corpus selection and comparability.

The most significant limitation of conducting this corpus analysis is the lack of a definitive interpretation of any phrase as generic. Naturally, other types of corpus studies are presented with a similar challenge, but it is particularly important to reiterate for an analysis of genericity, where ambiguity is prominent due to its inherent structural ambiguity. I will address alternative ways of textual analysis in the discussion in Section 2.8.2 below.

Bearing this in mind, this analysis is of an exploratory nature to illustrate how corpus analyses could contribute to the analyses of wide-spread linguistic constructions, which are not yet well-understood and neither studied making use of the whole linguistic toolkit. As with other types of linguistic analyses, this chapter will uncover its shortcomings and argue that ultimately, a combination of analyses and approaches is necessary to understand the construction as a whole.

2.8.1 Child-directed speech in English

The data in this section come from the Talkbank project (`https://talkbank.org`, MacWhinney, 2000), which hosts a collection of corpora that can be analysed via the CLAN program (Computerized Language ANalysis). To gain a better understanding of the input that children are exposed to during language acquisition, I use the CHILDES (Child Language Data Exchange System) database. This is a well-established, open-access database, which helps ensure that the data are comparable to previous research and aids efforts in reproducibility. Language interactions with children have been recorded in various regions of the world and in multiple languages on CHILDES. As Diessel (2009, p. 10) points out, the frequency of recordings "may provide enough information to trace the development of frequently occurring structures and expressions, [but] are not sufficient to study the development of less frequent phenomena". Due to the set-up of data collection in corpus analysis, a standard corpus only captures between 1.0-1.5% of linguistic input (Tomasello & Stahl, 2004).

I investigate the presence of generic subjects in child-directed speech in speakers of North American English, where the children are between the age of 0;6-5;1.[24] The two corpora chosen from the ones available on CHILDES were selected based on sample size, most commonly used child corpora (to aid comparability in this exploratory analysis), normally developing children with no known aphasia/language disorder, and naturalistic settings as opposed to directed lab sessions. These other types of corpora are also available on CHILDES but are specifically excluded here to achieve a complimentary picture of language acquisition and development, in addition to the experimental work designed for this thesis (although limitations will be addressed in Section 2.8.2).

Following these criteria[25], I have chosen to analyse Brent & Siskind (2001) (N=8, age range: 0;6-1;0; data collected in 1997), for analysis of linguistic input directed at preverbal children. For verbal interactions with young children before or just after the age of entering formal education, though they might include play in preschool environments, Brown (1973) (N=3, age range: 1;6-5;1; data collected between 1962-1966) was chosen.

The total number of children in the analysis equals 11. Since age is not always a reliable indicator of a child's linguistic capacities, searches for the mean length of utterance (MLU) were carried out on each individual .cha-file and summarised as a table for each participant's corpus. A sample can be found in Appendix A and

[24]This notation system expresses the 'exact age' of a child. This is calculated as: exact age = (*current date* - *birth date*)/365. I will use the age in this thesis as it is the system used in developmental psychology and translates into the recruitment criteria for the child-studies in Chapter 4.

[25]One could also use more recent corpora that are available to get an accurate picture of current language usage. However the Brown corpus, while more dated, is commonly used and a good point of comparison with other corpus studies. Initially, the analysis was to be run on a total of six corpora, with three corpora chosen to include children before entering formal education, and three having entered formal education at some point during data collection. These other four are Davis & MacNeilage (1995) (N=21, age range: 0;7-2;4; data collected between 1992-1994), Rollins (2003) (N=12, age range: 0;5-1;0; not clear when data were collected but necessarily between 1996-2003) for the former group and Sawyer (1997) (N=20, age range: 3;6-4;11; data collected between 1989-1992) and Gelman et al. (1998) (N=46, age range: 1;6-3;2; data collected mid-1990s) for the latter. This would render a sample size of n=110. Given the results of the analysis of the Brent & Siskind and the Brown corpus, I have decided that for the nature of this thesis, it provides sufficient insight into the benefits and limitations of such a study. These corpora were therefore not further analysed.

a summary of the averages for the Brent corpus in Table 2.1. The MLU between different children (files c1-w3) is similar, which is further reflected in the similarly small standard deviation of the child values, given as SD(CHI).

Manual checks were performed on all individual files. Most of them were accurate although sometimes the mother's tier (i.e. the mother's utterances themselves) were falsely attributed to the child or there were instances in which the audio track did not correspond to the transcript. In these cases, the specific .cha-file was marked up and excluded from further analysis rather than only taking out the utterances that were found to be falsely attributed. This way, it was not necessary to assume that the rest of the file was correct or that all mistakes were found manually. Minor errors were overlooked (e.g. a single, monosyllabic word that was misattributed to the child. Virtually no file was to be found completely free of transcription or other minor errors). The remaining recordings within one mother-child dyad were kept in the overall analysis.

Table 2.1: Summary of MLU averages in Brent & Siskind (2001)

CORPUS	MLU(CHI)	MLU(MOT)	SD(CHI)	SD(MOT)	#REC
Brent_c1	1.256	4.032	0.661	2.813	12
Brent_d1	1.525	4.261	0.376	2.957	11
Brent_f1	1.114	4.235	0.32	2.917	13
Brent_f2	1.15	3.022	0.204	2.334	11
Brent_j1	1.521	3.953	0.825	2.763	11
Brent_j2	1.007	3.056	0.031	2.462	14
Brent_m2	1	4.269	0	2.844	10
Brent_q1	1.05	3.335	0.15	2.814	7
Brent_s1	1.428	3.844	0.4	2.592	14
Brent_s2	1.123	3.81	0.223	2.554	13
Brent_s3	1	3.24	0	2.505	6
Brent_t1	1.084	4.205	0.166	2.844	11
Brent_v1	1.288	4.127	1.296	2.81	13
Brent_v2	1.026	3.871	0.153	2.793	11
Brent_w1	1.219	4.099	0.767	2.968	10
Brent_w3	1.059	3.959	0.219	3.005	11
AVERAGE	**1.178**	**3.832**	**0.362**	**2.748**	**11.125**

Note that for "number of recordings", the mother-to-child-ratio is not 1:1. This is because the children do not typically have utterances in every recording of the corpus associated with each mother-child-pair. Therefore, even if the number of recordings equals, e.g. 11, this does not necessarily mean that the averages calculated for the child are based on 11 recordings. Similarly, the number of recordings per family varied (for the Brent & Siskind, 2001 corpus between 6-14, with approximately 60-75 minutes of recorded interaction each) and the comparisons are therefore not based on equal number of hours recorded.

Both corpora were automatically tagged to allow for morphological analysis by adding a fully parsed tier (%mor). Each `.cha`-file was amended by running the superordinate command `mor *.cha` on them. This command runs `MOR, POST, POSTMORTEM`, and `MEGRASP` on the CHAT transcripts (see MacWhinney, 2000, note that this command also adds a %gra tier indicating the sentence's grammatical dependencies).[26] Consequently, the amended files allow for easier identification of strings in the corpora that are part of generic sentences by running morphosyntactically motivated commands instead of lexico-semantically-driven ones.

Table 2.2 provides an overview of the search terms performed, including the number of results for sample corpora:

The second search term specified in Table 2.2, which is the first of the more specific searches going beyond noun-verb combinations, looked for combinations of nouns (excluding pronouns and proper nouns) followed by a form of the verb "to be", with any intervening material:

(s2) `combo +s"N|*^*^cop|be&3S" +t%mor *.cha`

This is a catch-all for all phrases including an *NP+"is"* structure. In a corpus with child- or child-directed speech, this combination excludes a vast part of the data

[26]The full documentation of the output of the individual components of the `mor *.cha` command can be found in Part 3 of MacWinney, 2000 (in particular, Section 3 and 4). For current purposes, it is sufficient to know that first, the command adds information on the part-of-speech and grammatical categories linked to the CHAT transcripts. This additional information is stored in separate files associated with the lexical items of the corpus, in a separate folder named `lex`. Further, it adds information about metadata, such as punctuation. `MEGRASP` is short for "maximum entropy grammatical relation analysis of syntactic patterns", adding another tier to the existing %mor-tier, but should not be of relevance to the reader in this thesis.

Table 2.2: Overview of corpus analysis search terms and results for the Brown (1973) NB: all search terms started with "combo +s".

#	SEARCH TERM	GENERIC (POSSIBLY)
1	"N|*^*^v|*" +t%mor *.cha	0
2	"N|*^*^cop|be&3S" +t%mor *.cha	8
3	"det:art|the^*^cop|*" +t%mor *.cha	0
4	"det:art|the^*^cop|*^det:art|a" +t%mor *.cha	1 (4)
5	"kind^of" *.cha	461 (2)
6	"was^invented" *.cha	0
7	"was^discovered" *.cha	0
8	"was^introduced" *.cha	0
9	"introduce" *.cha	1
10	"invent" *.cha	0
11	"discover" *.cha	0
12	"died" *.cha	0
13	"extinct" *.cha	0
14	"is/are^common" *.cha	0
15	"is/are^rare" *.cha	0
16	"is/are^numerous" *.cha	0

as many interactions only consist of phrasal segments.[27] Even with the amended 3rd person form of *(to) be*, the copula-construction *is* still casts too wide a net as most of the results for this search term are episodic (e.g. "my[!] name is Ned." (000828.cha, line 2319), "teddy[!] bear is in the slinky" (000904.cha, line 4179)). Yet, there was still a small proportion of generic sentences that this search returned in the corpora of child-directed speech toward preverbal children. In Brent & Siskind (2001), there were eight utterances of a generalising nature, although only two of them can be categorised as having the form of a generic statement with a generic subject, i.e. the form of interest in the current research (where "*MOT" is the ID for the mother's tier):

(84) a. *MOT: medicine is purple (010224.cha, line 244)

b. *MOT: water is good for you (010221.cha, line 1641)

[27] Since one of the most common generic subjects is the BP, I initially ran a search where the form of *(to) be* was not restricted to the 3rd person singular. This search returned too many irrelevant results and the results were deemed to noisy.

As example (84-a) shows, one of them only has generic form but is not a true generalisation. Thus, in Brent & Siskind (2001), only one true generic statement can be found and the rest of the utterances with a generic flavour are questions that might prompt the child to think in general rather than specific terms. This is dependent on whether the child, without generic input, is capable of making the distinction between generic and specific reference at that point and even more so, making this distinction based on the grammar that the child might just be acquiring via the same input. Ultimately, this can be answered in terms of whether ambiguous statements are interpreted generically by default, i.e. whether the human conceptual system is organised under a *Generics-as-Default* system or whether genericity is learned through input, and likely after specific reference is acquired to tangible individuals in a child's environment. Note also that both of these subjects, 'medicine' and 'water', are mass terms, which are not the focus of the experimental work in this thesis (except for the mention of optional determiners in German in Section 2.2.1).

Another example that enables the child to extract generic knowledge also comes from the Brent & Siskind (2001) corpus (only quoting the mother's tier (MOT) for clarity, emphasis added):[28]

(85) MOT: oh these are some dumb birds, cupcake.
MOT: they came right up to it and didn't see it huh?
MOT: well.
MOT: **that bird (i)s a pigeon** (010111.cha, line 2689)

This statement can convey knowledge about a subkind, where it relates the subkind *pigeon* to the basic kind term *bird*.[29] It can also, like every generic subject, convey a specific interpretation, in this case referring to an instance of a pigeon. As a competent, adult speaker, this ambiguity is resolved via context. Example (85)

[28]Note that the %mor-tier annotates the demonstratives "these" and "that" and subject pronouns, but that no generic interpretation is referenced. However, in line with Bowdle & Ward, 1995, the demonstratives can act as generic demonstratives and are ambiguous. For *that bird (i)s a pigeon* in particular, the intended reference could easily be interpreted as "that kind of bird".

[29]Given the different semantic representations for kind and subkind statements identified in Chapter 2.3.3, I do not expect them to convey the same information, although they are intuitively more similar to each other than generic and episodic statements.

includes the immediately preceding context but even so, it remains ambiguous whether the mother refers to that instance of a bird or gives information about the subkind *pigeon* as "that (kind of) bird". Therefore, even though contextual information is available to the child, disambiguation requires an understanding of the fact that a distinction between generic and episodic reference can be made.

Interestingly, the IS in child-directed speech can perform a triple-function: First, it can be used to point to an individual instance, likely presented in a picture book that is read to the child, but also as e.g. *I baked a cake*. Second, it can identify a subkind, as in Example (85) above, where the *pigeon* is a subkind of the kind *bird*. Third, the IS can refer to a basic kind, e.g. *A bird is a feathered animal*. The IS is more commonly used than the DS, and even though the latter occurs with similar functions in the same circumstances, it only performs two out of the three functions (specific and kind reference).

Returning to the structural analysis listed in Table 2.2, further searches were carried out. In order to analyse the utterances in the corpus that are more likely to be generic phrases (I say "more likely" because there are always sentences that are ambiguous between an episodic and a generic reference), the search string is changed to a term that more specifically targets generic subjects, and is less likely to return instances of proper nouns. I altered the search commands to focus on DS subjects, to only include structures like "The X BE a Y":

(s3) `combo +s"det:art|the^*^cop|*^det:art|a" +t%mor *.cha`

This search yielded almost no results, with the few results that match the command typically not having generic reference. In the Brown corpus, this returned 21 results, none of which were clearly generic. Three cases were ambiguous, though a context analysis, executed manually, shows it is unlikely these sentences were intended generically. For example, the mother asked the child "Why are you going to do this?" and the child responding "The boy is a soldier." (Brown, 1973, 030707.cha, line 2581). The main problem with these instances is that the results cross sentence boundaries or have much intervening material that crosses clause boundaries, as the following example in Brown (1973, Mel's tier) illustrates (emphasis added):

> I think &-um **the** first time that I came (.) Kent called me to say that we **were** coming (.) or I guess it was **a** week we didn't come and he said (.).

To capture sentential fragments with generic reference that do not include a form of "to be", various searches were performed for lexical items or combinations thereof. Note first that "[t]here is no such thing as an exhaustive list of predicates forming the class of Carlsonian kind-level predicates(...) The paradigm case of a Carlsonian kind-level predicate is supposed to be *(be) extinct*, inclusive of all its variants: *is extinct, has died out, will become extinct soon, is in danger of extinction,* etc." (Mueller-Reichau, 2013, p. 76). Therefore, the results of these searches are to be interpreted as an indicator of the amount and the kind of generic input in child-directed speech, much like a corpus is a sample rather than a full representation of language itself.

The search for "kind of" in Brown (1973) rendered the most results out of all the searches so far.

(s5) `"kind^of" *.cha`

A manual inspection of the results yielded a total of 461 utterances that can be interpreted as making reference to a subkind (out of 527 hits total), with a further 2 being ambiguous. A parallel search in the Brent & Siskind (2001) corpus returned 48 occurrences. Two of those results were discarded as the transcript did not match the audio file. The remaining 46 results were split up into three categories: first, generic ones, for statements referring to subkinds; second, episodic ones for uses of "kind of" that refer to likeness; and third, ambiguous cases. Sentences with episodic (n=24) or ambiguous reference (n=6) were excluded from further analysis and left 16 results with a clearly generic reference. This means that only a third of the utterances including "kind of" likely refer to a subkind. Extending the search to include "kinda" returns an additional 84 occurrences. The results were sorted into the same three categories and show the following distribution: 62 results were episodic, 4 ambiguous, and 19 generic (in two instances, the utterances were part of the same phrase, thus a conservative estimate would be 17 generic phrases in the corpus). Similar to the "kind of" search, here just under a third (30.65% for 19 generic phrases for 62 results or 27.42%, counting 17 generics for 62 results) of utterances including "kinda" refer to a subkind. Finally, searching for "kinds of" returns 15 results in the Brent & Siskind (2001). As expected, since "kinds of" does not have the use of "likeness" that "kind of" and "kinda" have, almost all are clearly generic and only two are ambiguous since the transcription does not clearly indicate the NP following the expression, but it finishes with noise or inaudible monologue/dialogue (2/15 cases). They mainly co-occur with broad categories, rather than with lexical items that have a very specific use (see Table 2.3):

Table 2.3: Collocations of "kind of" in Brent & Siskind (2001)

COLLOCATIONS "kinds of"	FILE-LINE
different things	000910-1500
of stuff	001106-2796
things	001120-472
stuff	001120-1037
different things	001206-745
pieces	010104-100
plants and animals	010110-588
fruit trees	010110-657
food	010110-573
pretty things	010217-2875
stories	010224-262

Moving on from a search term that could indicate subkind reference, the next part of the analysis focuses on the aforementioned kind-level predicates, i.e. predicates that only apply to kinds. Searches (s6-s8) looked for *was invented*, *was discovered*, and *was introduced*,[30] none of which are applicable to an individual, in their function as kind-level predicates:

[30] Unlike "invented", both "discovered" and "introduced" have non-kind-level functions. An individual can discover that their phone has run out of battery, in which the predicate performs a function similar to "realising" or "noticing". Similarly, someone can be introduced to a new group of people, such that greetings are exchanged. Any findings of these uses of the predicates will not be taken into consideration for the analysis.

(s6-s8) `combo +s"was^invented/discovered/introduced" *.cha`

The search yields no results in Brown (1973) nor in Brent & Siskind (2001).[31] Broadening the search to look for instances of just the infinitive form, i.e. "introduce", "invent", "discover", (s9-s11) only returned results for "introduce" and none of the other two predicates, and none of them are in a generic sentence. Rather, their use is highly situational, as in "xxx think I invent these things now" (*MOT, 030320.cha, line 474). Including other predicates such as "died" or "extinct" (s12-s13) as a catch-all term for some variants of the Carlsonian kind-level predicates (see above, e.g. *is extinct, has died out, will become extinct soon, is in danger of extinction*) also does not return any results in the Brown corpus.

Even including predicates that are not strictly categorised as kind-level predicates (Mueller-Reichau, 2013, p. 77f), such as "(be) common, (be) numerous, or (be) rare" in the search does not change the overall picture significantly. The combinations in search (s14-16) returns no results.[32]:

(s14-s16) `combo +s"cop|*^adj|common/rare/numerous" +t%mor *.cha`

Less frequent predicates, such as "(be) eradicated", "(be) bred" or "(be) rampant" were henceforth excluded from the analysis, given the low occurrence of the most common kind-level predicates in the corpus.

Another point of interest raised by this corpus analysis is that all three subject types can be found in one utterance and used almost interchangeably. This becomes especially obvious when in the same situational context, the same kind of information is conveyed with different morphosyntactic structures. This change in the form of reference almost at random might facilitate a generic interpretation, given that these three morphosyntactic forms follow each other, as opposed to replacing the reference with a pronoun or demonstrative. The quote below from the Brent

[31] As explained in Chapter 2 above, Mueller-Reichau (2013) identified various types of kind-level predicates. In this corpus analysis, they have been group together since they do not render (distinct) results.

[32] I performed a pre-check in the MOR-folder for English to check for potential cross-categorisation of these adjectives and ran additional searches with the other categories. This did not affect the results and will therefore be disregarded in the main analysis.

& Siskind corpus exemplifies this particularly well (lines 2284-2452, edited for ease of legibility, all child-directed speech by the mother, where ‘xxx’ stands for unidentifiable speech, emphasis added):

> This is **a wolf**. Remember what **the wolf** says? There’s a wolf picture in the living room.
>
> This is a xxx. (...) **A nautilus**. This is. **The nautilus** is a xxx or xxx. Like an octipi [!] or a squid.

Under these circumstances, child-directed speech only names the animal with a singular subject phrase.

In summary, this exploratory corpus analysis has illustrated how the high occurrence of subjects that are potentially generic, either by form and/or by meaning, makes an analysis of the prevalence and distribution thereof difficult just by means of corpus analysis. Future research might focus more on semantically tagged corpora, although these are argued to be not accurate enough yet (see Reiter & Frank (2010) for an approach in computational linguistics, but also Declerk (1986) for a more general overview of the challenges due to various interpretations of generic sentences, even in conversational/pragmatic context). It has shown that recording naturalistic settings might not enable us to draw conclusions about the research question at hand to the extent that it can make a meaningful contribution to the literature. To address the limitations of finding generic subjects in naturalistic child-directed speech in a set of corpora, I will turn to a more targeted textual analysis of children’s books in the next section. This is based on the assumption that much of what children learn about the world is conveyed by others. As Gelman et al. (2013, p. 491) note so aptly, referring to observations made by Harris & Koenig (2006), “children would have difficulty learning about history, science, or religion solely by means of solitary interactions with the world around them”.

2.8.2 Limitations and discussion

The corpus analysis supports the assumption that directly kind-referring subjects, i.e. English DS subjects, are rare in child-directed speech. It also assumes that generic phrases with BP subjects are much more common. However, it is debatable when these instances are intended to be generic and even a liberal interpretation of genericity conducted on commonly assumed indicators of generic phrases yields limited results. This leads back to the initial question of whether, even if these phrases are intended as generic, they are interpreted by the child as such. Of course, the same argument can be made vice versa: Following the argument that generic interpretations are the default mode for the conceptual system, the child would equally struggle to interpret an ambiguous statement to make episodic reference.

The absence of kind-level predicates in the corpora analysed points to at least a puzzle as to how generic and kind-referring meanings are acquired for these ambiguous subjects. In terms of learnability, the rare linguistic input of kind-predicates needs to be accounted for. These are clear indicators of DS subjects being used in a kind-referring way and thus make the fact that young children seem to be able to understand kind-referring DS subjects at a young age even more surprising. An argument could be made in line with the general poverty-of-stimulus argument, i.e. that there are sufficient data points in the input and that the child is particularly capable of generalising from this limited exposure.

Alternatively, one could argue that searching for instances of generic statements in a vast corpus of child-directed speech is looking for a needle in a haystack (although, admittedly, the frequency of generic statements should far outweigh the frequency of needles in haystacks). Possibly, conducting a textual analysis of instructional child-directed speech might help us better understand the types of generic statements that children hear when learning about the world. While laboratory-based experiments in this thesis manipulate the role of morphosyntactic variation of the generic subject, we can start by taking an intermediate approach. Specifically, I will focus on the middle ground between a corpus analysis and a carefully designed experiment. I will present a brief textual analysis of a picture book for children about farm animals, which was used by Gelman et al. (1998), and of a picture book about trucks (selected based on previous research in Gelman et al., 2013, and personal communication with Susan Gelman).

For this analysis, not only is the distribution of different types of generic subjects important, but especially their co-occurrence. Children's books are a major part of how young children learn about the world in a Western context. While I acknowledge that this is not the norm, at the very least the child-directed speech in the corpora that form part of CHILDES target families of the same background. Thus, while this analysis might not provide us with the full picture of the acquisition of generics in English, it does not take away from the breadth of analysis a corpus study would provide us with.[33]

This analysis relies on the hypothesis that children acquire knowledge in both pedagogical and non-pedagogical contexts. Crucially, both children and adults seem to expect that generic knowledge is especially prevalent in pedagogical contexts, and adults even produce more generics in these settings as opposed to non-pedagogical contexts (Gelman et al., 2013).

Farm Animals by Helweg (1978) is a picture book for children that introduces them to various animals that a child could expect to live on a (North American) farm. The book was coded for generic phrases that had generic subjects by two coders. Intercoder agreement for generic subjects was 72%. All disagreements were resolved by discussion. Special consideration was given to the linguistic form of the subject as well as whether it was DEFINITE (variable: DEFINITENESS) and NUMBER, i.e. whether it was SINGULAR or PLURAL. Out of 67 phrases total, 52 were coded as having a generic subject, and a further 4 were coded as having a possibly generic subject. Looking at the subjects of the 52 definitively generic sentences, the highest component of them was made up of BPs (22), closely followed by pronouns referring back to the previously mentioned subject (16). A summary of the subjects is given in Table 2.4.

[33] In fact, families that are willing to participate, and thus give up their time, for scientific study might be more self-selecting and represent a more narrow demographic of the overall population than social interactions in which children are read picture books.

[34] Based on a breakdown of NUMBER, where 36 out of 52 generic DPs are plural, it is further possible to deduce that most pronouns are the plural 'they', as opposed to only very few singular pronouns, which were 'she' and 'it'.

Table 2.4: Subjects in the book *Farm Animals*

SUBJECT FORM	NUMBER OF OCCURRENCES
BP	22
IS	5
DS	6
pronoun	16[34]
DPl	2
mass	1

This shows that for animal kinds, children are likely exposed to BP subjects, or pronouns that refer back to categories previously introduced with a BP subject. In line with the corpus analysis, the acquisition of IS and DS subjects in their function as generic subjects seems surprising since their occurrence is relatively low. Of course, this is only one example but it aligns well with the findings from the CHILDES corpus data, which were exclusively analyses of natural and spontaneous caregiver-child interactions.

One limitation to consider in this context is that these types of books (and many of the books created and used for studies in this thesis) introduce animal categories. While the finer differences between animal categories and other types of categories will become more obvious in the later chapters, in particular the role of essentialism and tighter category-property links, I will briefly compare these findings with those for a book focusing on artefacts.

The Truck Book by McNaught (1978) follows a similar structure as *Farm Animals* but focuses exclusively on artefacts, and specifically different kinds of trucks. The book was coded for the same variables and intercoder agreement for generic subjects was 63%. All disagreements were resolved by discussion. Out of 75 phrases total, 48 were coded as having a generic subject, and a further 11 were coded as having a possibly generic subject. Compared to *Farm Animals*, this shower a higher level of ambiguity (indicated by a lower level of intercoder agreement) and a generally lower level of generic subjects. These could be attributed to the lower essentialism of artefacts. Looking at the subjects of the 48 definitively generic sentences, the highest component of them was made up of BPs (23), followed by IS subjects (10). A summary of the subjects is given in Table 2.5.

Table 2.5: Subjects in the book *The Truck Book*

SUBJECT FORM	NUMBER OF OCCURRENCES
BP	23
IS	10
DS	3
pronoun	6
DPl	0
mass	1
quantPL	5

This table shows that in child-directed, educational writing, the BP is the most common subject, regardless of the type of kind. However, within the story-telling context of the *Farm Animals* book, a higher proportion of pronouns was used, which might instruct children to understand genericity outside the most common generic subjects for count nouns in English. By contrast, *The Truck Book* has some quantified PL subjects which were interpreted as generic by both coders, something which is absent in the *Farm Animals* book, and not normally discussed in the linguistics literature on generics either. There is a higher ratio of IS subjects, which could be due to their subkind function seeing that *The Truck Book* introduces many types of trucks. As before, the ratio of DS subjects is still very low.

In summary, the corpus analysis and the picture books analyses demonstrate that singular generic subjects are rare in child-directed speech. This is in line with their more restricted usage but leaves open the question of how children are able to interpret these as generic from a young age on.

This chapter has provided an overview of theoretical approaches to the morphosyntax of generic subjects. These analyses have shown that the variation in distribution can be attributed to variation in the DPs. The following chapters take seriously how these morphosyntactic and semantic differences can lead to behavioural differences in experimental settings. These experiments explore the effect that morphosyntax has on the conceptual system, either alone or in combination with other cues. The next chapter will begin this enquiry by presenting a series of studies employing pseudowords to investigate the effect of morphosyntax while reducing the influence of world knowledge.

"Beware the Jabberwock, my son!
The jaws that bite, the claws that catch!
Beware the Jubjub bird, and shun
The frumious Bandersnatch!"

— Lewis Carroll (1871)

3

Pseudoword learning studies

Contents

3.1 (Why) Pseudowords for learning

Most of the knowledge that we have about the world is based not on direct observation but on knowledge passed on verbally by others (e.g. Gelman et al., 1998; Pelletier, 2010; Prasada et al., 2013, and somewhat more implicitly in Rips, 2011). This applies not only to the acquisition of world knowledge for young speakers during their childhood, but also to the continued acquisition of further (specialist) knowledge for adults.

In Chapter 2, we have already seen that the way we are able to communicate generic knowledge is limited by the content of the generalisations. In practice, this means that some statements, e.g. statistical ones, can only be expressed by a certain type of generic sentence. To recapitulate, I specifically address the question of why we judge a generic statement like *The tiger has stripes* to be grammatical but not **The car has a radio*, when both are acceptable with a bare plural (BP) subject: *Tigers have stripes* and *Cars have radios*. The generics literature has focused extensively on BP subjects, to a lesser extent on indefinite singular (IS) subjects (see Carlson, 1980; Lawler, 1973; Dahl, 1975), and to an even lesser extent on definite singular (DS) generic subjects. Interestingly, however, the BP is not always the preferred option. Some generic statements are easily expressed with DS subjects (*The dodo is extinct*), as contrasted with the BP (*Dodos are extinct*), and crucially disallow an IS subject (**A dodo is extinct*). Similarly, the IS is acceptable for sentences like *A tiger has stripes* but less so for sentences such as ?*A car has a radio*.

The aim of this chapter is to investigate in more detail the extent to which variation in the way we express generic knowledge can affect how we conceptualise the properties of categories. To reduce the effect of world knowledge as much as possible, and thus to isolate the role of morphosyntactic cues, the main means I have employed is the use of pseudowords in place of already existing, meaningful content words. Based on the hypothesis that these three generic subject forms have distinct semantic interpretations, the differences in their semantics are expected to lead to conceptual differences, observable in distinct behaviour. With careful design of the experimental stimuli, these studies should arrive at behavioural results that more closely reflect that there are differences in the semantics of generic subjects, rather than taking them as one joint class. Thereby, they can account for the sometimes-conflicting results in linguistics compared to psychology. This mainly concerns instances of limited acceptability of specific generic subjects, in the linguistics literature, compared to broader statements about the nature of genericity based on the (often exclusive) use of BP, in the psychology literature.

This chapter is structured as follows: I will begin by discussing theories of concept acquisition and the relationship between novel concepts and their properties in Section 3.2. In Section 3.3, I will start by reviewing established experimental methodologies that have previously addressed the interpretation of generics. In so doing, I will identify variables that can be manipulated in a more fine-grained fashion from a linguistic standpoint. This addresses concerns that the strong focus on BP subjects in previous experimental investigations of genericity has made these experiments insufficient to capture the intricate nature of generics. To account for the wide range of semantics of generics, I treat separately the three different types of generic subjects in English, indicated by their distinct linguistic make-up. Based on their morphosyntactic differences I incorporate them as distinct variables in the experimental design. In Sections 3.3.1-3.3.3, I will present findings from a series of learning studies that carefully explored the role of the individual morphosyntactic pieces of English generic subjects and the effect of training participants on knowledge they have already acquired implicitly. I then relate these findings to the prominent notion in psychology that generic interpretations are the default interpretation in the human conceptual system, i.e. *Generics-as-Default* (Gelman, 2004; Leslie, 2008; Meyer et al., 2011; Meyer & Gelman, 2016; Sutherland et al., 2015; Leslie & Gelman, 2012). While the interpretations of the semantic differences in this chapter do not rely on *Generics-as-Default*, the proposed analysis is not incompatible with this theory. Instead, I aim to add to the theoretical understanding of genericity by arguing that complex linguistic representations guide our conceptual system to hold various distinct generic interpretations. Sections 3.4-3.5 will refer back to notions of developmental theories introduced in Section 2.7. These studies will vary the amount of training that participants received on novel kinds, thus allowing us to analyse the effect of input. In Section 3.6, I will return to the variables NUMBER and DEFINITENESS and transfer the experimental methodology to create a version of these studies in French. I will compare the results of the French study to those from the studies in English. After summarising the findings, I will take into consideration the open questions that remain with regard to the role of morphosyntactic variation in the interpretation of generic phrases in Section 3.7.

3.2 Novel concepts and their properties

Novel words can function as a useful tool that approximates how we acquire knowledge, similar to the process that language learners undergo in their infancy, while enabling the testing of competent adult speakers of a language. These novel words label novel concepts, just as new terms are coined to refer to novel items, technology or phenomena in everyday life.

Recall that Section 2.5 provided an overview of the different types of properties that can be predicated of and attributed to a category. I focused particularly on the distinction between principled connections (PCs) and statistical connections (SCs) as two of the main types of properties. The experiments presented in the sections below will rely heavily on these notions. Prasada & Dillingham (2006, 2009) have named these *k-properties* and *t-properties*, based on whether they apply to the kind or the type, respectively. In line with the wider literature on concepts and genericity, I will continue using the terms *principled* and *statistical* properties and merely remind the reader of the parallelism by making it explicit again here. To recapitulate, PCs (or *k-properties*) are predicated of categories where the property is linked to the concept by an underlying, inherent mechanism. If the category lacks such a property, there is normally a cause that explains the property's absence. This concerns properties such as the four-leggedness of a cat or the egg-laying nature of ducks.

One way of testing whether a property is principally connected to a category is by using paraphrases, such as *by virtue of* or *is one aspect of* (as introduced by Prasada & Dillingham), in generic statements.[1] If the generic still holds, then a property is principally connected. Take, for example, the generic statement in (1-a) and compare it to the generic statements with these paraphrases in (1-b)-(1-c). All of these statements are equally acceptable:

(1) a. A cat has four legs.
 b. A cat, by virtue of being a cat, has four legs.
 c. Having four legs is one aspect of being a cat.

SCs, on the other hand, do not rely on an underlying mechanism that connects category and property. Statistically connected properties are only accidentally true of their categories, or true due to prevalence in the kind's members.

In a similar fashion to the paraphrases targeting PCs, there are paraphrases to test whether the category-property link is based on statistical connections. I will again draw upon those used by Prasada & Dillingham, which are *just happen to* and *just because most*. Like above, compare the generic statement in (2-a) with the paraphrases in (2-b)-(2-c):

(2) a. Cars have radios.
 b. Cars just happen to have radios.
 c. Cars have radios just because most cars do.

Crucially, these paraphrases do not overlap, such that the paraphrases for SCs do not render acceptable statements for PCs and vice versa, as illustrated in (3-a)-(3-b) and (4-a)-(4-b), respectively:

(3) a. #A cat just happens to have four legs.
 b. #A cat has four legs just because most cats do.

[1]Further, recall that Chapter 2 provides a semantic analysis of the differences between these two paraphrases. Ultimately, *by virtue of* and *is one aspect of* do not share the exact same underlying representation and seem to target slightly different aspects of generic statements. However, for the purpose of the studies to follow in this chapter, they provide a sufficient proxy for the distinction between PCs and SCs. This is not to say that future experiments should not think carefully about what the nature of PCs is and whether it is potentially a summary of multiple variables and should be seen as an umbrella category.

(4) a. #A car, by virtue of being a car, has a radio.

b. #Having a radio is one aspect of being a car.

In the experiments in the subsequent sections of this chapter, I make use of this distinction which I have argued affects the type of subject form that is licensed in English generic sentences. Before we turn to the studies themselves, remember that no property is inherently a principled or a statistical one. To illustrate, we know that *being striped* is a principled property of a zebra, and a statistical property of an index card. We know this because of our general knowledge of the nature of animal kinds and how they differ from artefacts.[2] With this in mind, the question these studies will address is whether the subject form of a generic sentence alone can, and does, elicit expectations as to how a property is connected to a category, in the absence of further knowledge about the category. The following sections address this question experimentally, through a series of learning studies, which also explore whether an extended training phase affects the process of linking a property to a specific type of subject.

3.3 English pseudoword studies: Novel critters with novel properties

The current set of learning studies are designed such that participants can develop a sense of how they think about unknown categories and their properties. Both the categories and properties are labelled with pseudowords, such that all content words are novel words expressing the new category and the only words in standard English are function words. To ensure that participants do not focus on memorising the novel words and try to recall them verbatim, the instructions explicitly mention that they should not focus on accurate memory retrieval.[3]

[2]Recall that Prasada & Dillingham (2006) categorise properties as *k-properties* and *t-properties* without indicating that any one property inherently falls within either one of these categories. For further discussion of this, refer also to Carlson (2010).

[3]While pseudowords are commonly used in psychological learning studies (e.g. Cimpian & Cadena, 2010), previous studies used either a novel subject or a novel property, but not both. Initially, it seems that using novel words in only one of the two crucial syntactic positions might facilitate the task. However, there are also some disadvantages to this design. For instance, when

Targeting PCs, I asked participants whether they think it is *one aspect of* the thing that it has the predicated property or whether they think the novel item has the property, *by virtue of* being the kind of thing it is (loosely based on Prasada & Dillingham, 2006, 2009). Targeting SCs, participants were asked whether they think the newly introduced object *just happens to* have the property or whether they think the object has the property *just because most* of these objects have that property.

3.3.1 Experiment 1 (def-num): BP vs IS subjects

Experiment 1 (DEF-NUM) compared the role of generic bare plural (BP) subjects, e.g. *Tigers have stripes*, with generic indefinite singular (IS) subjects, e.g. *A tiger has stripes*. In response to the questions in the testing phase, I hypothesised that participants would be more inclined to endorse statements about PCs, i.e. *by virtue of* and *is one aspect of*, when asked about the property that was introduced with IS subjects in the training phase. This is because PCs are asserted to be properties of each individual of a kind, unlike accidental or statistically prevalent properties. This expression of properties of an arbitrary individual is best expressed by an IS subject. I hypothesised that properties that were trained with BP subjects would be equally acceptable to express both PCs and SCs. This is because a bare, non-definite plural subject can express both the properties of an arbitrary individual and the properties of some individual that are perceived to be prevalent statistically due to the hypothesised semi-lattice structures of groups (as discussed in 2.3.2, with reference to Link, 1983 and Chierchia, 1998, inter alia). Alternatively, it is conceivable that IS subjects reduce the expectation of SCs because accidental information about individuals is not something that would be expected to be stored in the representation of a kind. To foreshadow the results, properties that

using novel subjects with existing properties, e.g. *Morseths have silver fur* (Cimpian & Cadena, 2010), participants often apply what they know about animals and how certain properties are linked to them (this idea will be further explored in Section 5.3 on overhypotheses). In this case, the conceptual connection between the *morseth* and the fur colour is expected to be principled, based on pre-existing general world knowledge of animals, and how the colour of their fur is determined by genetic mechanisms. Such recognition effects are less likely to occur with pseudoword-only stimuli. In the studies presented here, participants similarly learned via generic sentences, like *Krivels have a voove*. Afterwards, they were asked questions targeting the category-property links.

were trained with IS subjects were expected to be judged as more likely to be principally than statistically connected.

3.3.1.1 Materials

All stimuli are sentences attributing a property to a novel object. These sentences are composed of pseudowords in both subject and object position. Participants first underwent training and then moved on to the testing block with no delay in between. The training phase consisted of 30 sentences, each of which had one novel subject and one of two properties (*have a voove* or *have a larbial*). In the testing phase, participants saw 48 new sentences. These again had novel pseudoword subjects but occurred with one of the two previously introduced properties. The subject occurred in one of the two generic subject forms, i.e. the BP (e.g. *Morgots have a voove*) or the IS (e.g. *A teef has a larbial*). The category-property connection was not random. Rather, the subject was systematically connected with one of the two properties, such that the property always co-occurred either with a BP or an IS subject.

3.3.1.2 Participants

50 participants were recruited online via Amazon Mechanical Turk. All were adult native speakers of American English and received $2 in financial compensation.

3.3.1.3 Procedure

Participants first provided informed consent and filled out a short questionnaire about their background, collecting information on their age and first language. They then moved on to the training phase of the study. After agreeing to participate in the study, participants saw the following instructions:

> Welcome!
>
> You have arrived on a foreign island and are surrounded by many unfamiliar objects, creatures, and other things. Fortunately, you have met some researchers who will introduce you to some of the things around you, but these researchers also need your help in characterizing several new things too.

> First, they will tell you what they know about 30 things they have found on the island. The things they have found are unusual, so the researchers have made up a vocabulary to describe them.
>
> **Pay close attention to what they tell you as this will help you think about the characteristics of the new things that you and the researchers find.** You don't have to remember what specific things are like. Rather, try to get a general feeling of what the island may be like and what kinds of characteristics things can have.
>
> After this, the researchers will take you to a different part of the island. There, you will be asked what you think about the characteristics of the new things that you encounter.

During training, participants were introduced to 30 new objects with pseudoword names (*morgot, habin*, etc., in the BP or IS) that were paired with one of two novel properties, *voove* or *larbial.* Participants read the training sentences one at a time. The new properties were counterbalanced between participants, so that they were not all exposed to the same subject-property combination, thereby preventing potential phonological or semantic associations from distorting the results. Recall that for each participant, one property was always presented with the same subject type. For example, if the property *voove* co-occurred with the BP (5), the property *larbial* was presented with an IS subject (6):

(5) Morgots have a voove.

(6) A habin has a larbial.

At the end of the training phase, participants moved on to the testing phase and were given the following instructions:

> Good job!
>
> That was the first part. We will now move on to a different region of the island that the researchers haven't explored yet. They hope that the characteristics they have given you so far will be useful in helping them answer some questions about the new things they find there. The researchers will ask you to answer some questions about the things they encounter based on what you have learned so far.
>
> You may not always feel very certain about your answer, so try to go with your initial reaction to the question.

Participants were then presented with 48 new pseudowords in subject position, each of which was combined with one of the previously introduced properties (*voove, larbial*). Participants first saw one of four prompts of the following forms: 'Excellent! We've found a pimsette and it has a voove', 'Oh look! That's a cleeg, and it has a voove', 'Here's a cleckle and it has a voove', 'We thought we might find a moff and this one has a voove'. To keep participants engaged and prevent them from forming response strategies, there was random variation in these introductory exclamatives across the two previously introduced properties.

Participants answered questions designed to probe their understanding of the connection between a new kind and the trained properties. The answers about this new object provided information to determine whether they assumed the property to be principally or statistically connected. For each testing sentence, participants were asked to rate one of four questions, adapted from Prasada & Dillingham (2006, 2009). PCs were tested with questions of the type *Do you think that a cleeg, in virtue of being a cleeg, has a voove?* or *Do you think that it is an aspect of being a cleeg that it has a voove?* Statistical connections were targeted with questions that took the form *Do you think that cleegs just happen to have a voove?* or *Do you think that cleegs have a voove just because most cleegs have a voove?*[4]

3.3.1.4 Analysis

Responses were given on a 1-7 Likert scale with 1 = "Definitely No" and 7 = "Definitely Yes". Response times were collected to ensure participants carried out the task appropriately.

The data from three participants were excluded due to very fast response times. The results are based on the responses of the remaining 47 participants.

[4]There was another set of questions to test participants' understanding of the connection between concept and property. Each of the questions presented here also existed as a negated version, e.g. *Do you think that cleegs do not just happen to have a voove?* or *Do you think that it is not an aspect of being a cleeg that it has a voove?* However, these questions were not only complicated and ambiguous but also behaved differently than the rest of the questions. Furthermore, this thesis is not specifically interested in negation and the intricacies that embedded negated questions entail before understanding the relationship between morphosyntax and properties under investigation first. Therefore, I do not include them in the report and analysis, and neither were these questions included in any follow-up experiments.

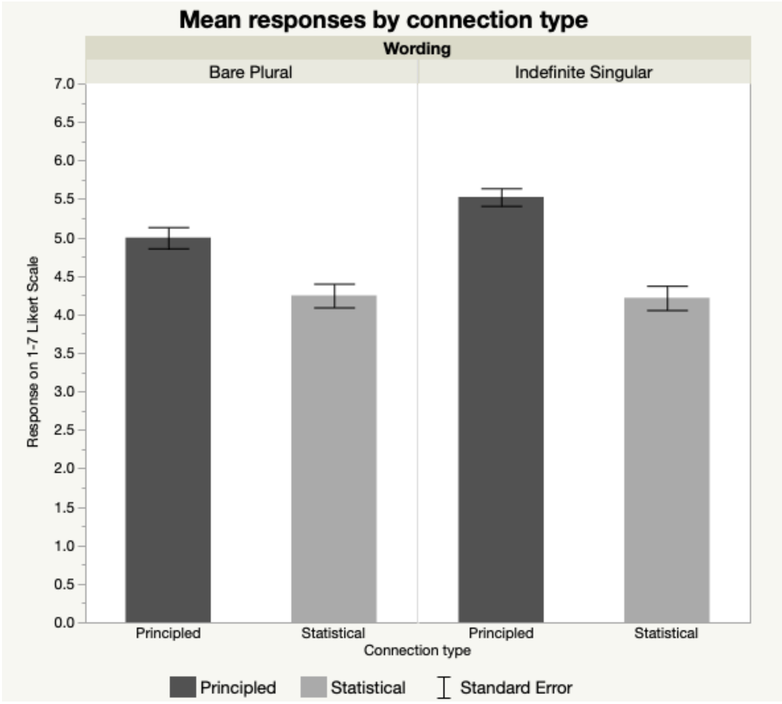


Figure 3.1: Comparison of responses given for questions targeting principled and statistical connection, by subject form with which they were trained. Error Bars = ± 1 SE

An analysis of the remaining response time data found no significant differences between conditions.

A 2 (wording: BP vs IS, within participants) × 2 (connection type: PC vs SC, within participants) ANOVA was conducted over subject-averaged ratings (by item) from a 1-7 Likert Scale to establish whether the differences between the ratings for generic subject types for each of the property types were significant. This analysis revealed a main effect of connection type ($F(1,722) = 52.4452$, $p < .0001$), but no main effect of wording ($F(1,722) = 3.0436$, $p = .0815$). There was a significant interaction between wording and connection type ($F(3,722) = 3.8572$, $p = .0499$). These effects are plotted in Figure 3.1.

Planned comparisons (Tukey HSD) show that participants were more likely to endorse questions targeting PCs during the testing phase if that property had been trained with IS subjects compared to BP subjects ($p = .0458$). Participants were not significantly more or less likely to endorse an SC question based on the wording ($p = .9986$)

3.3.1.5 Results and discussion

The results (a significant interaction of wording and connection type, indicating higher ratings for PCs if trained with IS subjects) suggest that participants made use of the grammatical context in which a new property was given to later judge the conceptual connection that property was likely to have with a new object. For IS subjects, participants appeared to be sensitive to its denotation of atomic individuals, and when this type of subject is used generically, they interpret the property to be one that applies to each individual. Properties that apply to each individual generically are likely to be principally connected, and therefore the IS leads participants to encode the property as likely to be principally connected and generalise its use accordingly.

The results from this study provide support for the hypothesis that IS subjects raise expectations that a connection between subject and property is principled. Experiment 2 (DEF-NUM), presented in the following section, was designed to follow up on this and to investigate whether DS subjects equally raise these expectations of PCs between kind and property.

3.3.2 Experiment 2 (def-num): BP vs DS subjects

Experiment 2 (DEF-NUM) directly extends Experiment 1 (DEF-NUM) by comparing BP subjects with definite singular ones, e.g. *The triangle has three sides.* Recall that in Chapter 2.3, I have provided a reanalysis for the DS as a definite NUMBER-less subject (DN). The use of 'DS' in this chapter should be interpreted as equally referential to 'DN' subjects. Based on acceptability judgements observed for English kind-referring statements, I predict that properties that are trained with DS subjects should also enhance the endorsement of statements targeting PCs. However, as in Experiment 1 (DEF-NUM), it is conceivable that DS subjects alternatively reduce SCs because accidental information about individuals is not something we expect to be stored in the representation of a kind.

3.3.2.1 Materials

For the comparison between BP and DS subjects, the material also consisted of 30 training and 48 testing sentences and the subject was again systematically connected with one of the two properties, such that the property was either primed with a BP or DS subject. The sentences were of the form *Morgots are a voove thing* or *The teef is a larbial thing.* The form of the sentences was changed according to a {SUBJECT [BE] a PROPERTY thing} template to ensure high accessibility of generic rather than episodic readings.[5] The paraphrase *in virtue of* was changed to *by virtue* for this and all subsequent *pseudoword* studies.

3.3.2.2 Participants

48 new participants were recruited online via Amazon Mechanical Turk under the same criteria and conditions.

3.3.2.3 Procedure

The procedure was similar to the one used in Experiment 1 (DEF-NUM). During the training phase, participants were introduced to 30 new objects with pseudoword names (*morgot, habin*), which were either presented as BP or as DS subjects. The properties were linked to either one of the morphosyntactic subject forms (counterbalanced between participants), such that if the property *voove* co-occurred with the BP, participants saw (7). In this case, the property *larbial* was presented in the DS (8):

(7) Morgots are a voove thing.

(8) The habin is a larbial thing.

[5]This was done due to concerns that individual reference is more likely for DS generics than IS ones, based on their more restricted use. As Chapters 4 and 5 will show, this is actually not the case.

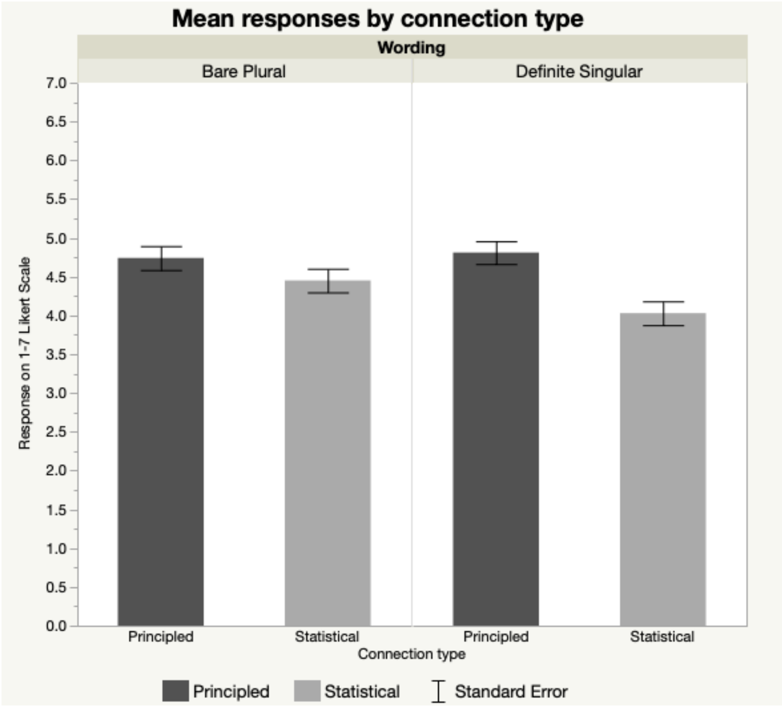


Figure 3.2: Comparison of responses given for questions targeting PCs and SCs, by subject form with which they were trained (BP and DS). Error Bars = ± 1 SE

3.3.2.4 Analysis

The data from two participants were excluded due to very fast response times. The results are based on the responses of the remaining 46 participants. The analysis of the remaining response times again found no significant differences between conditions.

A 2 (wording: BP vs DS, within participants) × 2 (connection type: PC vs SC, within participants) ANOVA was conducted over subject-averaged ratings (by item) from a 1-7 Likert Scale to establish whether the differences between the ratings for generic subject types for each of the property types were significant. This analysis revealed a main effect of connection type ($F(1,707) = 12.4327$, $p = .0004$), but no main effect of wording ($F(1,707) = 1.3384$, $p = .2477$). The analysis did not find an interaction between wording and connection type ($F(3,707) = 2.6138$, $p = .1064$). These effects are plotted in Figure 3.2.

Due to a lack of interaction in the omnibus model, no paired comparisons were carried out.

3.3.2.5 Results and discussion

The results of this study (main effect of connection type no interaction of wording and connection type) suggest that the grammatical context in which a property is presented may be significant. It may affect how participants judge the conceptual connection that a property is likely to have with a new object. For cases involving DSs, I hypothesised that participants use their generic interpretation as kinds to interpret the property to be one that applies to kinds. However, DS subjects behave unlike IS subjects: First, there was no interaction between DS subjects and property type. While IS subjects enhance expectations of principled connections compared to BP subjects, it is conceivable that DS subjects might diminish expectations of SCs. This effect, however, was not significant in this study. A follow-up study will investigate whether it is possible to show this interaction. The proposed hypothesis for this is that DS subjects behave asymmetrically to IS subjects and lower participants' expectations that a property is statistically connected to the kind. This is because properties that apply to kinds are unlikely to be merely statistically connected.To better understand participants' responses, and to consider an improved design going forward (to address the marginal significance yet reliable trends seen in the results of these first experiments), the next section presents a replication of Experiment 1 (DEF-NUM) and Experiment 2 (DEF-NUM).

3.3.3 Replication with subset of stimuli

Given that only one of the first two studies showed the hypothesised interaction, the items were analysed and a subset of them was re-run. As noted in Footnote 4 above, the questions that included a negation did not return any consistent data patterns. Addressing the effect of negated, embedded questions is not the main focus of this research and we therefore excluded these questions in the replication of Experiment 1 (DEF-NUM) and Experiment 2 (DEF-NUM), i.e. Experiment 1-R (DEF-NUM) and Experiment 2-R (DEF-NUM).

3.3.3.1 Materials

The stimuli are the same as those analysed and presented in Experiment 1 (DEF-NUM) and Experiment 2 (DEF-NUM). That means that the training phase is unchanged, but that no negated sentences were presented to participants during the testing phase.

3.3.3.2 Participants

50 new participants were recruited for each Experiment 1-R (DEF-NUM) and Experiment 2-R (DEF-NUM).

3.3.3.3 Procedure

The procedure was the same as in Experiment 1 (DEF-NUM) and Experiment 2 (DEF-NUM). Note that participants were still asked to respond to 48 test questions but that none of them had negated paraphrases. Therefore, the analysis of responses includes twice the number of responses than Experiment 1 (DEF-NUM).

3.3.3.4 Analysis of Experiment 1-R (def-num)

The data of all 50 participants were analysed. The analysis of the response times again found no significant differences between conditions.

A 2 (wording: BP vs IS, within participants) × 2 (connection type: PC vs SC, within participants) ANOVA was conducted over subject-averaged ratings (by item) from a 1-7 Likert Scale to establish whether the differences between the ratings for generic subject types for each of the property types were significant. This analysis revealed a main effect of connection type ($F(1,2335) = 36.9254$, $p < .0001$), but no main effect of wording ($F(1,2335) = 1.7902$, $p = .181$). The analysis did not find an interaction between wording and connection type ($F(3,2335) = .1495$, $p = .699$). These effects are plotted in Figure 3.3.

Due to a lack of interaction in the omnibus model, no paired comparisons were carried out.

For Experiment 1-R (DEF-NUM), comparing BP vs IS subjects, the same general patterns that were observed in Experiment 1 (DEF-NUM) were replicated, but the interaction between wording and property type was, crucially, not replicated.

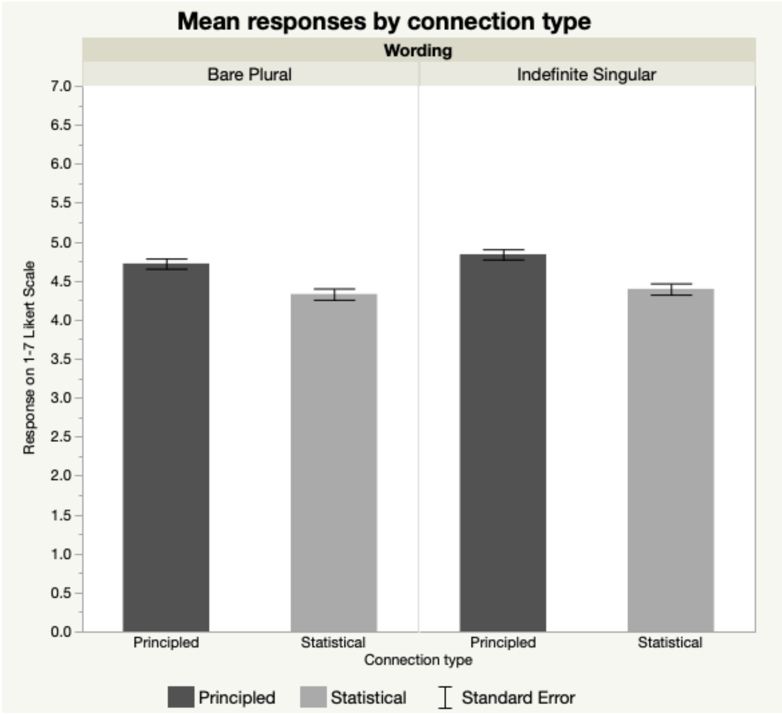


Figure 3.3: Experiment 1-R (DEF-NUM): Mean scores given for questions targeting PCs and SCs, by subject form (BP vs IS). Error Bars = ± 1 SE

Note that generally, the statistical paraphrases (i.e. *just happen to*, *just because most*) have a wider range of response values. Figure 3.4 shows how both for BP subjects and for IS subjects, the bottom of the interquartile range is lower even though the means are similar across all four bars. Most notably, the responses for PCs within IS subjects provide the most homogenous response pattern. The lower quartile is at 4 (out of 7 on the Likert-scale). This smaller spread provides further evidence for participants' inclination to agree more with PC paraphrases if trained with IS subjects.

The next section will present the analysis of the replication of Experiment 2 (DEF-NUM).

3.3.3.5 Analysis of Experiment 2-R (def-num)

The data of five participants were excluded as they were either repeat participants or had very fast reaction times. The analysis is based on the remaining data of 45 participants. An analysis of the remaining response times again found no significant differences between conditions.

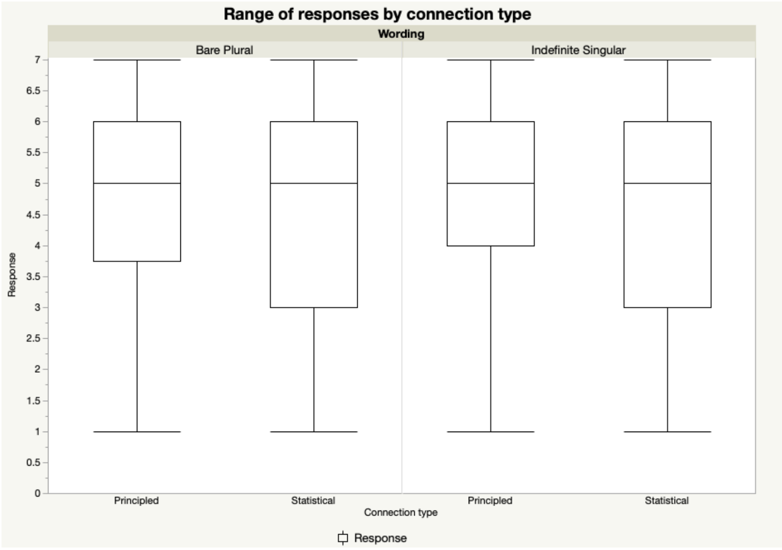


Figure 3.4: Response range for BP and IS subjects based on responses to questions targeting PCs and SCs

A 2 (wording: BP vs DS, within participants) × 2 (connection type: PC vs SC, within participants) ANOVA was conducted over subject-averaged ratings (by item) from a 1-7 Likert Scale to establish whether the differences between the ratings for generic subject types for each of the property types were significant. This analysis revealed a main effect of connection type ($F(1,2145) = 62.4861$, $p < .0001$), but no main effect of wording ($F(1,2145) = .1423$, $p = .706$). The analysis did not find an interaction between wording and connection type ($F(3,2145) = 2.3439$, $p = .1259$). These effects are plotted in Figure 3.5.

Due to a lack of interaction in the omnibus model, no paired comparisons were carried out.

Similar to Experiment 1-R (DEF-NUM), the statistical paraphrases (i.e. *just happen to*, *just because most*) have a wider range of response values. Figure 3.6 shows that within SCs, both BP and DS subjects have lower values for their lower quartile (with the lower quantile dropping to 3 out of 7 on the Likert-Scale), even though the means are comparable across all four conditions. By comparison, within PCs, both BP and DS subjects have a smaller range of response values. This smaller spread provides further evidence for participants' inclination to agree more with PC paraphrases, in this case regardless of whether trained with BP or DS subjects.

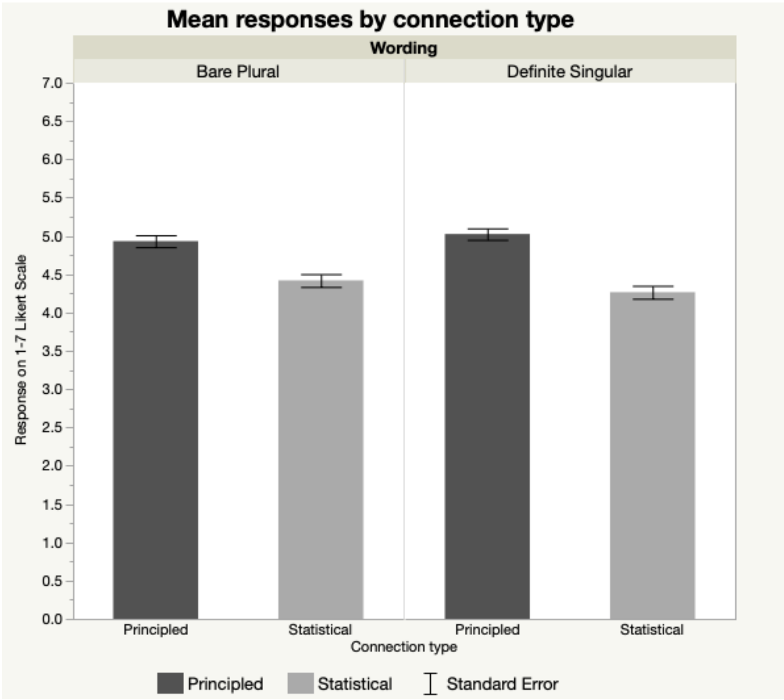


Figure 3.5: Experiment 2-R (DEF-NUM): Mean scores given for questions targeting PCs and SCs, by subject form (BP vs DS). Error Bars = ± 1 SE

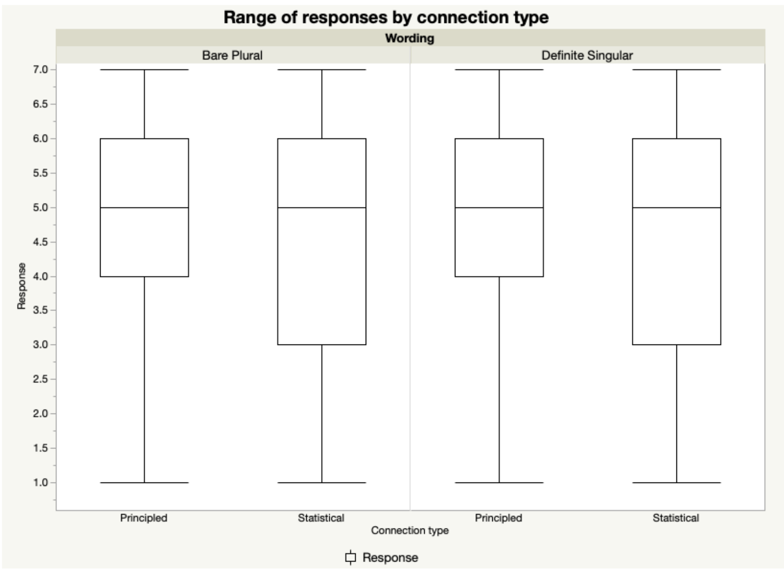


Figure 3.6: Response range for BP and DS subjects based on responses to questions targeting PCs and SCs

3.3.3.6 Results and discussion

The findings of the replication studies (main effect of connection type) support the previous findings that participants are more likely to agree with questions that include PC paraphrases compared to SC paraphrases. Both of the initial studies (Experiment 1 (DEF-NUM) and Experiment 2 (DEF-NUM)) were re-run without the stimuli that included negations, and both of these replications, Experiment 1-R (DEF-NUM) and Experiment 2-R (DEF-NUM), show a main effect of connection type, with PCs rated higher than SCs.

In general, the relative response patterns were the same between Experiment 1 (DEF-NUM) and Experiment 1-R (DEF-NUM), as well as between Experiment 2 (DEF-NUM) and Experiment 2-R (DEF-NUM). Only the absolute values differed, such that the ratings were higher across the board in Experiment 1 (DEF-NUM) compared to Experiment 1-R (DEF-NUM). The reverse was the case for the absolute values of Experiment 2 (DEF-NUM) and Experiment 2-R (DEF-NUM). Again, this held across the board. All of these means are summarised in 3.1.

Table 3.1: Mean ratings given by participants to questions targeting PCs and SCs on a 1-7 Likert scale for Experiment 1 (DEF-NUM) - Experiment 2-R (DEF-NUM)

Experiment	BP-PC	BP-SC	IS-PC	IS-SC	DS-PC	DS-SC
1	5.00	4.25	5.53	4.22	-	-
1-R	4.72	4.33	4.84	4.39	-	-
2	4.74	4.45	-	-	4.81	4.03
2-R	4.93	4.42	-	-	5.02	4.27

Comparing the ratings of connection types across the wording conditions, Experiment 1 (DEF-NUM) showed that PCs were rated higher when trained with IS, compared to BP. These results were also found in the replication studies. Asymmetrically, the ratings of SCs were lower when the property was trained with DS, as opposed to BP subjects. This finding was also confirmed in the replication experiments.

However, these studies do not provide reliable support for the hypothesis that, in the absence of other information, the conceptual system relies on information conveyed by the morphosyntax of the subject to propose a likely type of category-property link. In other words, the morphosyntactic form of a generic subject did not consistently show that participants are more likely to rate questions targeting PCs higher if they were trained with an IS subject. Neither the results show that participants are more likely to rate questions targeting SCs lower if they were trained with a DS subject. Only Experiment 1 (DEF-NUM) showed the hypothesised interaction.

In the next section, I will address the questions of learnability and innateness by looking at the effect that the training phase has on how participants conceptualise these category-property links. It is possible to hypothesise that this knowledge about the relationship between morphosyntax and connection type is already present in participants, but that the separate training phase is not conducive to tapping into this knowledge. Therefore, I present two follow-up studies, each of which varied the amount of training participants received before responding to the questions about the novel kinds and properties.

3.4 Tapping into existing morphosyntactic knowledge

The results of the first set of studies might be subject to scepticism regarding learnability, or more specifically the role of the training phase. One might argue that if the type of morphosyntactic knowledge that these experiments tap into is something that speakers readily employ, they should not need to be trained on the combination of one of the two properties and a particular morphosyntactic subject form. In this case, participants should not need a separate training phase in which they familiarise themselves with the morphosyntactic form of a subject that it has when it co-occurs with a specific property. Thus, participants could simply be shown a sentence with a pseudoword subject and a pseudoword property, and immediately respond to a question that targets their understanding of the type of connection that they expect between the subject and property.

Alternatively, one might argue that the amount of new information requires repeated input in order to understand the conceptual information conveyed. The BP subject in particular is ambiguous in that it may express either a principled or a statistical connection and might not lead to straight-forward answers without a minimal amount of training. Both of these hypotheses are explored in the rest of this chapter. Section 3.4.1 starts by exploring the first hypothesis, i.e. that the training has already taken place in competent speakers and should not be necessary. Section 3.5 then looks at the effect that more training has on the strength of the connection between a novel property and novel categories.

3.4.1 Experiment 3 (def-num): BP vs IS vs DS subjects (no training)

In this first follow-up study, participants did not undergo a separate training phase. It is possible to hypothesise that participants already use their pre-existing knowledge of the relationship between morphosyntax and the types of connections they may express. Thus, they could see a novel generic statement and immediately be tested on their understanding, directly after being exposed to the generic sentence. If participants rely on their pre-existing morphosyntactic knowledge, they should behave similar to the previous studies. This is because under this hypothesis, the training phase does not provide the participants with any information that is relevant or useful to answering the questions in the testing phase.

Alternatively, other processes might influence their judgements, in particular the role of cognitive load. The related processes of seeing (and potentially trying to remember) novel categories and properties, even if it is explicitly stated that this is not the task of the experiment, might be taxing for the participant in a way that prevents them from answering the questions in a coherent way.

To summarise, the hypothesis for this no-training condition is that, if participants use their pre-existing knowledge of the relationship between morphosyntax and the types of connections they may express, their performance should be similar to the experiments presented in Section 3.3. On the other hand, the training phase could help participants familiarise themselves with the type of task and methodology, and enable them to access their judgements in a series of questions that are separate from this training phase.

3.4.1.1 Materials

The stimuli are similar to those used in the previous studies, with two key differences. First, the was no separation into a training and a testing phase. Instead, the presentation of the training sentence was directly followed by one of the four questions that target participants' expectations as to the type of connection. To illustrate, participants were presented with a generic statement, e.g. *Krivels have a larbial* or *A mospeth has a voove.* On the same screen, they were asked the same questions introduced above that target principled or statistical connections, e.g. *Do you think krivels have a larbial just because most krivels do?* or *Do you think a mospeth, by virtue of being a mospeth, has a voove?*

Second, BP, IS, and DS subjects are now compared within the same experiment, meaning that each participant is exposed to all three subject types during the study. Previously, the comparison of BP subjects with IS and DS subjects was separate (split into Experiment 1 (DEF-NUM) and Experiment 2 (DEF-NUM)). The assignment of properties to specific morphosyntactic subject forms was maintained, such that this study has a 3 × 2 within participants design (subject form: BP, IS, DS × connection type: PC, SC). The experiment consisted of 30 pairs of sentences and questions for each condition, resulting in a total of 120 questions per participant.

3.4.1.2 Participants

33 participants were recruited on Amazon Mechanical Turk under the same criteria as in the previous experiments. None of them had participated in any of the previous studies.

3.4.1.3 Procedure

After consenting to participate and filling in a short sociolinguistic questionnaire, participants started the study. The overall procedure was similar to the one in Experiment 1 (DEF-NUM)-Experiment 2-R (DEF-NUM), with one significant difference. Participants did not undergo a separate training phase in this study. Instead, they saw the generic statements for 30 novel objects with novel properties but then immediately answered a question about this statement directly afterwards. The "training" and testing sentence were presented on the same screen. To illustrate, a full trial took the form of

Rampres have a wogerp.

Do you think that rampres just happen to have a wogerp?

Each new pseudoword subject was combined with a new property. There were three pseudoword properties, such that each of them was presented with one of the three subject forms. As in the previous studies, the connection was not random but counterbalanced between participants. Each property always co-occurred with the same subject form. The questions were the same questions that were used in the previous experiments, targeting participants' understanding of the connection between kind and property. They again responded on a 1-7 Likert Scale, where 1 = "Definitely No" and 7 = "Definitely Yes". Two paraphrases targeted PCs (*by virtue of, is one aspect of*) and two targeted SCs (*just happen to, just because most*).

3.4.1.4 Analysis

Two participants were excluded due to very fast response times. The results are based on the remaining 31 participants. An analysis of the response times found no significant differences between conditions.

A 3 (wording: BP vs IS vs DS, within participants) × 2 (connection type: PC vs SC, within participants) ANOVA was conducted over subject-averaged ratings (by item) from a 1-7 Likert Scale to establish whether the differences between the ratings for generic subject types for each of the property types were significant. This analysis revealed a main effect of connection type ($F(1,3651) = 175.6655$, p

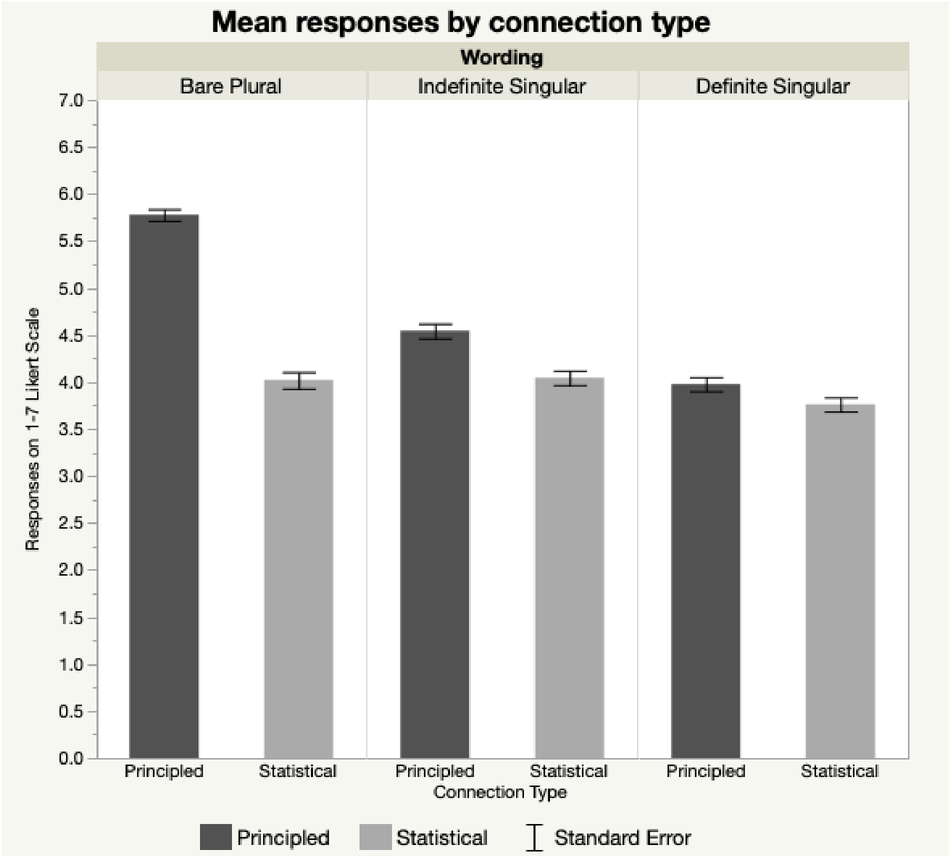


Figure 3.7: Experiment 3 (DEF-NUM): Mean scores given for questions targeting PCs and SCs, by wording (BP vs IS vs DS). Error Bars = ± 1 SE

$< .0001$), and a main effect of wording ($F(2,3651) = 92.4912$, $p < .0001$). The analysis also found an interaction between wording and connection type ($F(5,3651) = 58.3156$, $p < .0001$). These effects are plotted in Figure 3.7.

Planned comparisons (Tukey HSD) show that participants were more likely to endorse a PC question if that property was presented with a BP compared to an IS subject ($p < .0001$) or a DS subject ($p < .0001$). Participants were also more likely to endorse a PC question if that property was presented with an IS subject compared to a DS subject ($p < .0001$). For SC paraphrases, there were no differences in participants' endorsements based on the wording with which the generic statement was presented (all p-values between .0928 - .9999).

3.4.1.5 Results and discussion

The main effect of connection type in these experiments was driven by the effect on trials in which a BP subject was presented. However, within PC trials, all three generic subject types elicited significantly different responses. This means that each type of generic subject produced different levels of expectations in participants regarding how the (novel) property is assumed to be connected to the novel category described by the generic statement. Surprisingly, the subject form that rendered the highest ratings for questions targeting PCs was BP, followed by the IS, and finally the DS. The BP rated higher than either of the singular generic subjects is opposite the predictions I made for these studies. This might indicate that without any training, participants prefer BP subjects and that only with more exposure to the less common generic subjects are these also rated higher for questions targeting PCs.

Note that the findings from Experiment 1 (DEF-NUM), which showed an interaction of between connection type and IS subjects and raised expectations that a property was principally connected to the category, were not replicated. To explain the lack of an interaction for IS subjects, I hypothesise that participants need a certain amount of exposure to novel properties within generic subjects. This was previously provided via the training section in this learning paradigm. If they do not receive such a training period, they might be subject to a cognitive overload. This might be more important for IS subjects as they are less commonly used in everyday generalisations. Further, the general pattern for BP subjects shows that participants are at least attuned to the effect of generic language in the absence of other clues. Thus, the same mechanisms might be employed as in the original experiments, but seeing that there was no separate training phase, participants did not have as much time to familiarise themselves with the properties and how they were linked to the generic subjects.

In order to test the hypothesis that training is a necessary part of conceptualising category-property links, the following experiment will expose participants to more training before asking them to judge the connection type.

3.5 Providing more training

3.5.1 (Arrested?) Development

Before presenting the next set of pseudoword studies, which employed a paradigm with an extensive training phase in contrast with the no-training condition in Section 3.4.1, I will briefly recapitulate the relevant assumptions about linguistic and conceptual development first introduced in Chapter 2.7. One of the main assumptions of the experiments presented so far in this chapter is the notion that pseudoword studies emulate the developmental process of acquiring generic knowledge. The studies in this section specifically homes in on this idea. Following Prasada's well-known (2010) observation that our knowledge of categories exists despite limited direct personal experience of the instances of these categories, I will manipulate the amount of training to better understand its role and significance. The role of exposure has been discussed in the developmental literature, especially seeing that children learn about the world, and do so at an impressive rate.

One might argue that more exposure to information about a category, or, in this case, to a property, will lead to a richer mental representation of the property's structure. Under this line of argument, the link between a category and its properties continually strengthens based on repeated input. This is a particularly prominent argument in prototype theories. Rips (2011, p. 181f) presents a range of theories of perceptual categorising (citing Posner & Keele, 1968, 1970; Reed, 1972) and summarises: "if people have to classify, for example, schematic faces into two previously identified sets, they mentally compute a prototype for each set, where the prototype specifies the average values of the members of that set..." Note that this is a process that assumes visual stimuli and a task that will result in two separate categories. More to the point of this thesis, this process relies on categorisation of properties that are statistically prevalent. Based on the arguments explored in Chapter 2, we know that generics are exception-tolerant, have variable truth-conditions, and often express non-obvious properties of categories. As we will see later on in Chapter 7, there might in fact be parallel processes to categorisation which need not rely on repeated exposure.

This leads us to an alternative prediction: it is not unreasonable to assume that the adult participants recruited for this experiment have already passed through a phase of extensive learning and possess the implicit knowledge that specific subject forms imply a specific type of property that is either principally or merely statistical connected to the concept. In this case, any extra learning would be of limited benefit, as the learning stage is already completed. This is also a worthwhile hypothesis in light of the findings of the no-training studies.[6]

In the sections below, I will present experiments designed to test the hypothesis that learning can help strengthen category-property links. This relies on the assumption that learning is grounded, at least to some extent, on repeated exposure. In this case, it is the repeated presentation of a specific property with a specific generic subject form, which is hypothesised to underline the type of connection that a property has to its category, by virtue of its form. The experiments below therefore employ an extensive training paradigm, based on the studies presented above.

3.5.2 Experiment 4 (def-num): BP vs IS subjects (extensive training)

Based on the findings in Experiment 1 (DEF-NUM) through Experiment 3 (DEF-NUM), one could argue that participants do in fact tap into their morphosyntactic knowledge for these learning studies and make use of the relation between morphosyntactic form and connection type. However, it seems that there was an effect of the training in Experiment 1 (DEF-NUM) and Experiment 2 (DEF-NUM) that ultimately affected how participants responded to the questions, compared to the no-training condition in Experiment 3 (DEF-NUM).

[6] Additionally, one might argue that the learning in the training phase is different to the initial learning in language acquisition. Rather than learning that there exists a relation between wording and connection type in general, participants in these studies 'learn' to apply this to novel properties without the additional benefit of having world knowledge. This learning relies more strictly on the observation that the BP allows for both PCs and SCs, while the IS and DS are more restricted, and consequently more appropriate to express certain types of connections than the more ambiguous generic BP.

Therefore, a follow-up study was designed to test the effects of extending the amount of training participants receive prior to the testing phase. While Experiment 3 (DEF-NUM), without a separate learning phase, did not fully replicate the initial findings (although it did show a main effect of connection type), extensive training might enhance these effects, including the initially observed interactions.

The hypothesis for Experiment 4 (DEF-NUM) is that participants' responses will show the same pattern as in Experiment 1 (DEF-NUM) and Experiment 2 (DEF-NUM), and potentially at a greater magnitude. However, it is conceivable that more domain-general mechanisms might be at work, which could overload the participants by exposing them to even more novel creatures. Thus, the extensive training could lead to the same cognitive demands as in Experiment 3 (DEF-NUM).

3.5.2.1 Materials

The materials used in these experiments are similar to those used in the previous experiments, with the exception that we included more novel kinds. We trained participants on twice the number of training items (60 training sentences). The number of testing items remained at 48. In this study, we compared BP and IS subjects, keeping the stimuli in line with the initial two experiments that only employed a two-way distinction.

3.5.2.2 Participants

49 new participants were recruited on Amazon Mechanical Turk under the same criteria as in the previous experiments.

3.5.2.3 Procedure

The overall procedure was the same as for Experiment 1 (DEF-NUM) through Experiment 2-R (DEF-NUM).

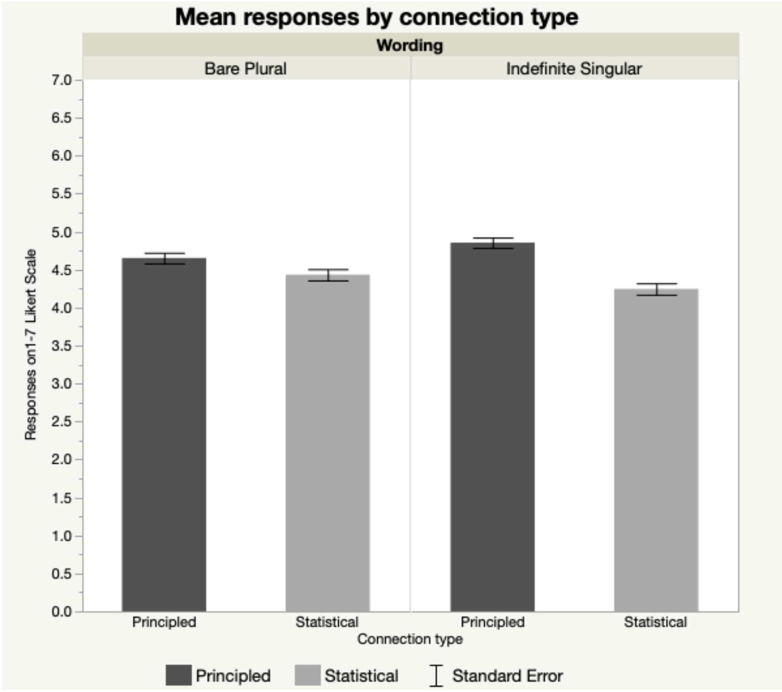


Figure 3.8: Comparison of means for BP vs IS by connection type. Error Bars = ± 1 SE

3.5.2.4 Analysis

The data of two participants were excluded due to very slow reaction times or missing trials. The analysis is based on the remaining data of 47 participants. An analysis of the response times found no significant differences between conditions.

A 2 (wording: BP vs IS, within participants) × 2 (connection type: PC vs SC, within participants) ANOVA was conducted over subject-averaged ratings (by item) from a 1-7 Likert Scale to establish whether the differences between the ratings for generic subject types for each of the property types were significant. This analysis revealed a main effect of connection type ($F(1,2218) = 33.2342$, $p < .0001$), but no main effect of wording ($F(1,2218) = .0107$, $p = .9176$). The analysis also found an interaction between wording and connection type ($F(3,2218) = 7.3377$, $p = .0068$). These effects are plotted in Figure 3.8.

Planned comparisons (Tukey HSD) show that participants were not more likely to endorse a question targeting PCs during the testing phase if that property had been trained with IS subjects compared to BP subjects ($p = .1925$). Neither

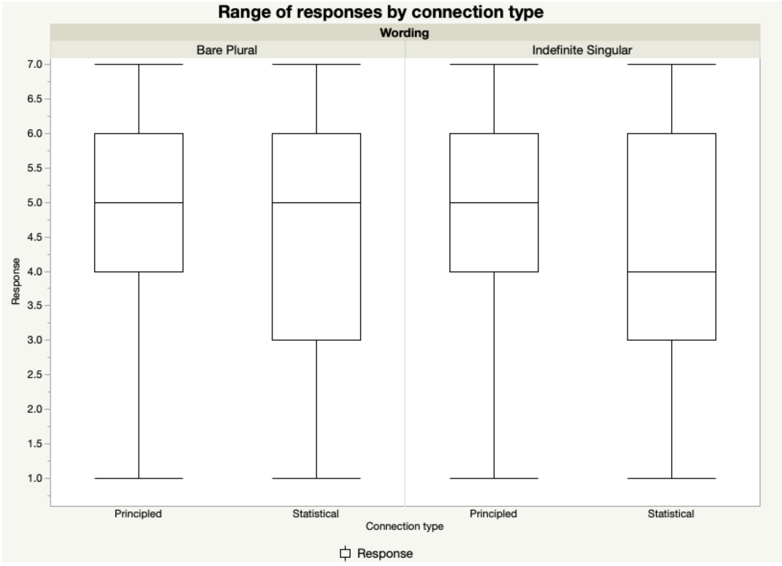


Figure 3.9: Response range for BP vs IS by connection type

were participants more or less likely to endorse a question targeting SCs based on the wording (p = .2538).

Similarly to the previous experiments, the statistical paraphrases elicited a wider range of response values. Figure 3.9 shows that for both BP and IS subjects, the bottom of the interquartile range is lower for questions targeting SCs. The lower quartile is at 3 (out of 7 on the Likert-Scale). One notable difference in this dataset is that the median for questions targeting SC paraphrases, if they were trained with IS subjects are relatively lower than all the other conditions. Within PCs, both BP and IS subjects elicited a smaller range of response values, as was the case in the other experiments.

3.5.2.5 Results and discussion

Overall, the interaction of wording and connection type observed in Experiment 1 (DEF-NUM), eliciting higher ratings for questions targeting PCs for properties trained with IS subjects, was also found in this version of the pseudoword studies with more extensive training. There were no differences between the ratings for questions targeting PCs or SCs if these were trained with BP subjects.

One possible explanation for finding this interaction with more training, but not reliably finding it with a medium amount of training, and neither in the no-training condition, could lie in the role of the amount of training that participants receive. It seems that in general, participants need some amount of training to learn about the novel properties and their co-occurrence patterns with the generic subjects. This is supported by the comparatively weaker results in the no-training condition. However, before coming to overall conclusions on the effect of extensive training, the next section will compare BP subjects to DS ones, such that the paradigm is extended to this kind-referring subject form. It is possible that the rarely used generic DS subject did not show interactions in any of the previous studies because it requires extensive training.

3.5.3 Experiment 5 (def-num): BP vs DS subjects (extensive training)

3.5.3.1 Materials

The stimuli are the same as in Experiment 4 (DEF-NUM), with the exception that the IS subjects were replaced by DS subjects.

3.5.3.2 Participants

53 new participants were recruited via the online participant recruitment and research platform Prolific under the same criteria as in the previous experiments.

3.5.3.3 Procedure

The procedure was the same as in the previous experiments.

3.5.3.4 Analysis

The data of three participants were excluded due to very fast reaction times. The analysis is based on the remaining data of 50 participants. An analysis of the remaining response times found no significant differences.

A 2 (wording: BP vs DS, within participants) × 2 (connection type: PC vs SC, within participants) ANOVA was conducted over subject-averaged ratings (by item) from a 1-7 Likert Scale to establish whether the differences between the ratings for generic subject types for each of the property types were significant. This analysis

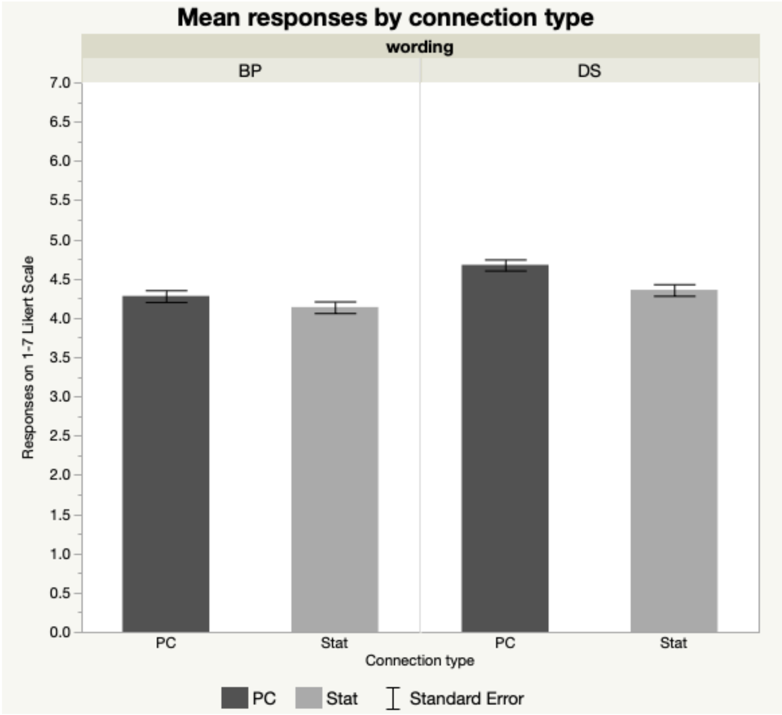


Figure 3.10: Experiment 5 (DEF-NUM): Mean scores of BP and DS by connection type. Error Bars = ± 1 SE

revealed a main effect of connection type ($F(1,2396) = 9.8095$, $p = .0018$), and a main effect of wording ($F(1,2396) = 17.6291$, $p < .0001$). The analysis did not find an interaction between wording and connection type ($F(3,2396) = 1.4469$, $p = .2291$). These effects are plotted in Figure 3.10.

Due to a lack of interaction in the omnibus model, no paired comparisons were carried out.

Figure 3.11 shows that the range of responses is similar across all four conditions. However, the range of responses showed that the median of DS responses, regardless of whether tested by questions targeting PCs or SCs was higher than the median for BP.

3.5.3.5 Results and discussion

In this extensive training version of the pseudoword study paradigm, participants were inclined to agree more with statements targeting PCs if these were trained with DS subjects. This differs from the findings of Experiment 2 (DEF-NUM) and Experiment 2-R (DEF-NUM), where paired comparisons did not find differences in

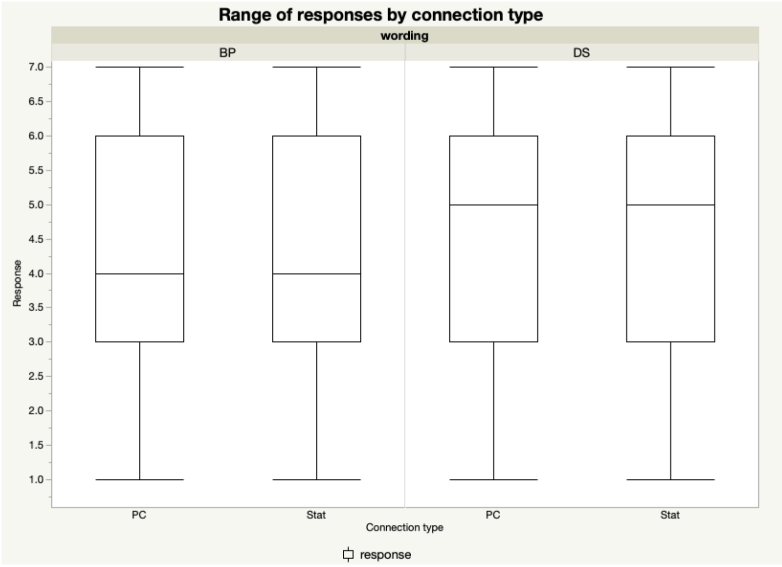


Figure 3.11: Range of responses for PCs and SCs, by wording (BP vs DS)

endorsement within questions targeting PCs. Interestingly, this is also the opposite pattern to that found in the no-training condition in Experiment 3 (DEF-NUM). In line with developmental considerations, as well as a generally more frequent use of BP generic subjects in everyday speech compared to DS subjects, the data suggest that with more training, participants are increasingly more likely to endorse statements that express PCs when they are presented with a DS generic subject.

There was, however, no interaction between property type and DS subjects. Thus the results are still mixed and do not clearly indicate whether DS subjects lead participants to provide lower ratings for questions targeting SCs. The only reliable effect that these studies have shown is that PCs are more easily generalised than SCs, i.e. a main effect of connection type in all of the pseudoword studies in this chapter.

Returning to points raised in the discussion of the previous experiment, Experiment 4 (DEF-NUM), it seems unlikely that participants paid less attention in the extended training phase. The difference in results could also be explained by the fact that DS generic subjects are relatively rare and the extended training particularly benefits this subject type, whereas less training is necessary for IS generic subjects in a pseudoword paradigm.

I will move on to some general comparisons of all *pseudoword* studies in the next sections, summarising the main findings of all of the experiments in this chapter so far.

3.5.4 Interim conclusions

In all of the learning studies presented in this chapter, participants made judgements without knowing what kind of category the generic statements referred to. Therefore, any effects relating to category-property links can only be attributed to the role of morphosyntax. This is unlike other studies where participants either make inferences based on visual stimuli or where information about category membership is provided in the instructions. Initially, one might assume that this design makes a generic interpretation less accessible overall, but the studies found generally high agreement with statements, in line with other studies on generic statements, e.g. those that compare generic to quantified or specific statements (Gelman et al., 2010; Leslie et al., 2011; Prasada et al., 2013, among others).

All of the pseudoword studies consistently showed a main effect of connection type, regardless of the morphosyntax of the subject. This supports the hypothesis that, in general, people are more inclined to accept generic statements that indicate, by virtue of their paraphrase, that properties are principally connected to a category. An overview of the various analyses is presented in Table 3.2.

Table 3.2: Comparison of main effects and interaction of wording and connection type for all *pseudoword* experiments

	Exp 1	**Exp 1-R**	**Exp 2**	**Exp 2-R**	**Exp 3**	**Exp 4**	**Exp 5**
Wording	n.s.	n.s.	n.s.	n.s.	*	n.s.	*
Connection Type	*	*	*	*	*	*	*
Wording × Connection Type	*	n.s.	n.s.	n.s.	*	*	n.s.

I have also noted that the mean ratings were higher across the board for the replication studies Experiment 1-R (DEF-NUM) and Experiment 2-R (DEF-NUM), where the negated questions were excluded from the testing phase. However, this observation does not take into account the interactions that were crucial for the initial hypothesis that the morphosyntax of the subject affects the interpretation of the connection type between concept and property. This interaction was only found in Experiment 1 (DEF-NUM), Experiment 3 (DEF-NUM), and Experiment 5 (DEF-NUM).

Table 3.3 first shows that, in general, participants gave more positive responses to questions targeting PCs than those targeting SCs. This was consistently shown as a main effect of connection type in all studies. Within properties that were trained with BP subjects, questions targeting PCs elicited more positive responses than those targeting SCs (with grand mean response values of 4.88 and 4.31, respectively). Turning to the properties that were trained with IS subjects, we find that more positive responses were given for the questions targeting PCs (grand mean of 4.94), while questions targeting SCs in these experiments were also rated relatively low (with a grand mean of 4.24). Finally, properties trained with DS subjects received relatively high ratings for questions targeting PCs than for those targeting SCs (grand means of 4.61 and 4.09, respectively). However, following the hypothesis that SCs are less appropriate to be expressed by generic DS subjects, we should focus more on the asymmetrically lower ratings for SCs in the DS wording condition. Firstly, the DS-SC condition has the lowest grand mean of all conditions. Secondly, the DS-PC condition is still rated higher than the IS-SC and BP-SC conditions, reflecting a general generic interpretation of this subject type.

Based on these means, the no-training condition seems to heavily favour positive judgements for PCs trained with BPs (mean = 5.81). The means in this condition lower with an increased amount of training (Experiment 4 (DEF-NUM), Experiment 5 (DEF-NUM)). Inversely, the more training participants receive on DS subjects, the higher the ratings. This seems to mirror developmental trajectories in which DS subjects are initially less frequently used in generic contexts. Further experiments in Chapter 5 will more specifically test generic interpretations of novel kinds in adult and child participants.

The overall means are visually plotted in Figure 3.12.

Post-hoc paired comparisons across conditions revealed that there were cases in which participants differed in their level of agreement with statements expressing PCs. The pattern that emerged showed that with more exposure, participants

Table 3.3: Response values on a 1-7 scale of connection type by wording for all *pseudoword* experiments.
BP = bare plural subject, IS = indefinite singular, DS = definite singular. PC = principled connection, SC = statistical connection.

Exp	BP-PC	BP-SC	IS-PC	IS-SC	DS-PC	DS-SC
1	5.00	4.25	5.53	4.22	-	-
1-R	4.72	4.33	4.84	4.39	-	-
2	4.74	4.45	-	-	4.81	4.03
2-R	4.93	4.42	-	-	5.02	4.27
3	5.81	4.13	4.54	4.08	3.91	3.71
4	4.66	4.43	4.83	4.26	-	-
5	4.28	4.14	-	-	4.68	4.36
Grand Mean	4.88	4.31	4.94	4.24	4.61	4.09

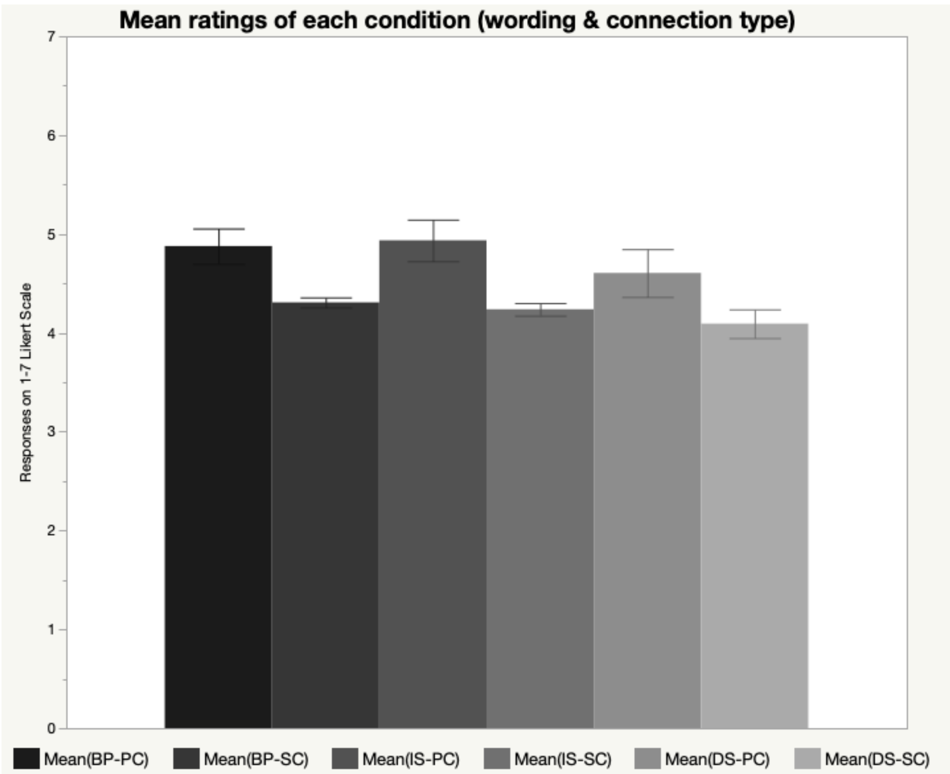


Figure 3.12: Mean ratings given on 1-7 Likert scale of connection type & wording for all *pseudoword* studies. Error Bars = ± 1 SE

were more likely to endorse statements expressing PCs if the property was trained with a singular generic (IS or DS). In cases where participants had received no separate training phase, the opposite pattern holds and participants were more likely to endorse questions targeting PCs if they were trained with a BP. None of the comparisons of statements expressing SCs showed any preference in subject type, regardless of the amount of training. These patterns are summarised in Table 3.4.

Table 3.4: Relative preferences, indicating participants' agreement with statements targeting PCs and SCs, depending on the wording of the subject form, in paired comparisons for all *pseudoword* experiments

Exp	PC	SC
1	BP < IS	BP = IS
1-R	BP = IS	BP = IS
2	BP = DS	BP = DS
2-R	BP = DS	BP = DS
3	BP > IS > DS	BP = IS = DS
4	BP = IS	BP = IS
5	BP < DS	BP = DS

Thus, these studies provide only limited support for the idea that one specific subject form can influence how properties are generalised, and could ultimately account for relative preferences between generic subjects.

I will present one follow up to further investigate whether the conceptual system relies on morphosyntactic cues, in the absence of other information. If this is truly a hallmark of the conceptual system, these effects, and possibly the hypothesised interactions that were only sometimes found, should be found in other languages as well. The next section will explore the same question under the same experimental paradigm but in French, keeping in mind and taking advantage of the distinct ways in which this language expresses generic subjects.

3.6 French pseudoword studies: The definite plural

3.6.1 Cross-linguistic differences in morphosyntactic cues

So far, we have focused on the variation in the morphosyntax of English generic subjects for count nouns. Before describing how the pseudoword paradigm can be applied to the French language, I will recall briefly some broader points on the relationship of conveying semantic differences via morphosyntactic cues from a cross-linguistic perspective. Chapter 2 addressed the linguistic variables that are involved in computing the distinct meanings of subject DPs. Specifically, I have focused on NUMBER and DEFINITENESS.

The morphosyntactic distribution of generic subjects found in English is, of course, only a small subset of how languages may express the various underlying semantic structures that generic subjects can represent. For English, this chapter and the preceding one have argued for distinct subject forms for direct kind reference (the DS), atomic instances of kinds (the IS), and groups of atoms (the BP). Other languages may use different combinations of NUMBER and DEFINITENESS to achieve the same or different semantic representations. Chapter 2 has presented examples from Armenian, Yucatec Maya, Jakaltek Maya, German, Spanish, Brazilian Portuguese, and others, which all use various combinations of these means and others, e.g. classifiers, to express genericity. This variation in morphosyntactic cues, which should allow languages to arrive at similar representations via different means, will be useful for experimental means. Using different subject types, hypothesised to carry out similar functions, enables a comparison of the behavioural results in the same paradigm under the same conditions.

This section will focus on French, which, as a Romance language, has a different distribution of determiners (and therefore DEFINITENESS). Examples (9)-(10) provide an overview of the distribution of French generic and non-generic subjects (these examples are adapted from Lee, 1996). In French generic subjects, determiners are obligatory (see examples (9-a), (9-d), and (9-e)), and no generic reference can be made with a bare plural subject (9-b) and (9-c). Of the available generic subjects, the most common one is the definite plural (DPl) (as in (9-a)), which is seen as the default option. Thus, intuitively, the DPl in French seems to perform the same function as the BP in English. Another interesting case seems to be (9-f), with an indefinite plural subject, which looks as if it provides subkind references.[7]

(9) a. Les oiseaux volent.
The bird.PL fly.PL
'Birds fly.'

b. *Oiseaux volent.
Bird.PL fly.PL
intended: 'Birds fly.'

c. *Oiseau vole.
Bird flies.
intended: 'Birds fly.' / 'The bird flies.'

d. Un oiseau vole.
A bird flies.
'Birds fly.' / 'A bird flies.'

e. L'oiseau vole.
The.bird flies.
'Birds fly.' / 'The bird flies.'

f. #Des oiseaux volent.
Indef-PL bird.PL fly.PL
'Birds fly' (often: some bird species flies)[8]

The distribution of generic subjects for French mass nouns (as in (10) below) mirrors more closely the distribution found in English. Most notably, the range of expressions is more limited than for count nouns (recall the covert structures of the DP discussed in Chapter 2). The only acceptable subject type for generic reference with mass nouns is the definite singular (10-a). Bare nouns (10-b) and nouns with the indefinite determiner 'de' (10-c) are judged unacceptable. While

[7]I thank my friend Célia Souque for judgements, clarifications, and patience on all matters French generic and quantified statements.

[8]This still assumes that all normal members of the kind, and therefore implicitly a majority, fly. To overtly express that *only some* birds fly, this would better be expressed by 'Certains oiseaux volent.' I was further told that this is normally contextually licensed, emphasising the aspect that *some* (kind of) bird flies. Crucially, individual reference is not licensed (although it is unclear whether it might be possible to coerce it).

the indefinite determiner ‘un(e)’ is acceptable for count nouns, it does not allow generic reference for mass nouns (10-d).

(10) a. L’eau est transparente.
The.water is transparent.
‘Water is transparent.’

b. *Eau est transparente.
Water is transparent.
intended: ‘Water is transparent.’

c. *D’eau est transparente.
Indef-SG.water is transparent.
intended: ‘Water is transparent.’

d. #Une eau est transparante.
A water is transparent.
‘One water is transparent.’ (but not water in general)

To compare the effects observed in the English pseudoword studies, the final part of this chapter will present the results of a version of the above presented pseudoword experiments in French.

The reason the stimuli from the studies presented earlier in this chapter have been used for a French version in Section 3.6.2 is because of their easy comparability and similarity in the make-up of generic subjects. French offers a close comparison with English generic subjects, while differing in one of the crucial variables manipulated in the studies in Section 3.3.

3.6.2 Experiment 6 (def-num): DPl vs IS subjects

I hypothesise that the learning of pseudoword properties in combination with pseudoword subjects, based solely on co-occurrence of morphosyntactic information, is a feature of our conceptual system. As such, I expect it to apply not only in English, but cross-linguistically. Whether the property occurs with a DPl or an IS subject should aid the conceptual system of the participants in making judgements about the connection type between novel category and novel property. The following sections explain how the stimuli from the original studies have been adapted to French and present the results from this extension.

3.6.2.1 Materials

The materials used for the English version of the studies were translated into French. The pseudowords were taken from the French Lexicon Project (Ferrand, 2010) and adapted to adhere to our linguistic requirements.

Unlike English, French nouns have one of two genders: masculine or feminine. For the design of these stimuli, a gender variable was taken into consideration as well and all pseudoword nominals were assigned a gender: one of the novel properties was feminine and the other was masculine (*une sépite, un vitin*). Likewise, the novel categories were evenly split into masculine and feminine gender (e.g. *un rouchant, une distelle*).[9]

The questions (including paraphrases) were translated by a French native speaker to ensure the same intuitions would apply. For PCs, the paraphrase *is one aspect of* was translated to "est une caracteristique pour", while *by virtue of* was translated to "du fair qu'il soit". This renders stimuli such as "Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une nible?" ("Do you think having a sépite is one aspect of being a nible?") and "Est-ce que tu penses qu'un déri, du fait qu'il soit un déri, a un vitin?" ("Do you think that a déri, by virtue of being a déri, has a vitin?"). Targeting SCs, the paraphrase *just because most* was translated to "simplement parce que la plupart des" and the paraphrase *just happen to* was translated to "comme par hasard". To illustrate, this renders questions like "Est-ce que tu penses que les pleutrons ont une sépite simplement parce que la plupart des pleutrons ont une sépite?" ("Do you think that pleutrons have a sépite, just because most pleutrons have a sépite?") and "Est-ce que tu penses que les sersalets ont, comme par hasard, un vitin?" ("Do you think that sersalets just happen to have a vitin?"). The co-occurrence of the properties with either DPls or IS subjects was counterbalanced between participants.

[9] I thank Daniel Cornford for help with the creation of these stimuli.

3.6.2.2 Participants

74 subjects were recruited via the online survey platform Prolific. The experiment itself was hosted on Ibex-farm.

3.6.2.3 Procedure

The procedure was the same as in the experiments conducted in English. After filling in the consent form and a sociolinguistic questionnaire, participants saw a translated version of the introduction as above (all stimuli are in Appendix B.2). Participants underwent a training phase of 30 novel objects and then saw the break page, telling them they would move on to a novel part of the island. As in the English versions, they then saw 48 new objects presented with the same two properties they were trained on (*sépite, vitin*). They were introduced to the question with a prompt of the following or similar form: *Regarde! Ça c'est un voilot et il a une sépite.* ("Look! That's a voilot and it has a sépite."). These exclamatory prompts were randomly varied and followed by one of the four questions targeting the participant's understanding of the connection between kind and property. Responses were given on a 1-7 Likert Scale.

3.6.2.4 Analysis

The data of four participants were excluded due to their self-reported status of not being a native speaker of French.[10] There were no significant differences in reaction time between conditions.

A 2 (wording: DPl vs IS, within participants) × 2 (connection type: PC vs SC, within participants) ANOVA was conducted over subject-averaged ratings (by item) from a 1-7 Likert Scale to establish whether the differences between the ratings for generic subject types for each of the property types were significant. This analysis revealed no main effect of connection type ($F(1,3356) = 1.892$, $p = .1691$), but a

[10] To account for slow reaction times, in addition to the analyses presented in the main body of the text, all analyses were also carried out excluding data points where reaction times to respond to the question were higher than 10 secs for that trial. In cases where this happened for more than 10% of all trials for any specific participant, the participant was excluded. Even though this led to a significant data loss, the results were unchanged.

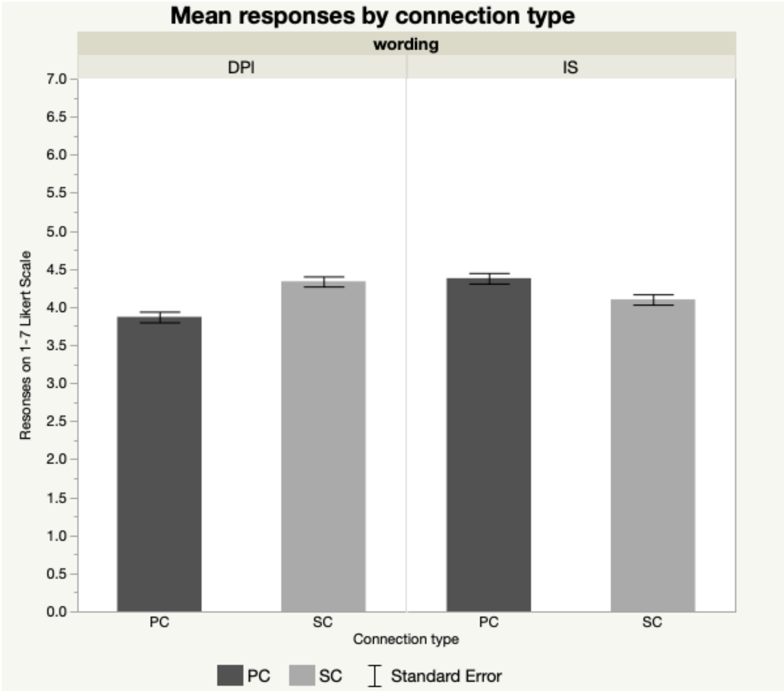


Figure 3.13: Mean scores by connection type. Errors bars are ± 1 SE from the mean.

main effect of wording (F(1,3356) = 3.9397, p = .0472). The analysis also found an interaction between wording and connection type (F(3,3356) = 29.6993, p < .0001). These effects are plotted in Figure 3.13.

Planned comparisons (Tukey HSD) show that participants were more likely to endorse a question that targeted PCs during the testing phase if that property had been trained with IS subjects compared to DPl subjects (p < .0001). Participants were marginally more likely to endorse a question targeting SCs if that property had been trained with DPl subjects compared to IS subjects (p = .0682).

This asymmetry is, to an extent, reflected in the range of responses. While the English studies had a wider spread of responses for questions targeting SCs as opposed to PCs, the French data show that the spread of responses is larger for SCs if trained with IS subjects and PCs if trained with DPl subjects, as plotted in Figure 3.14. As in the previous studies, a smaller spread provides support for the hypothesis that participants are more likely to endorse statements targeting SCs, if trained with DPls and statements targeting PCs, if trained with IS subjects.

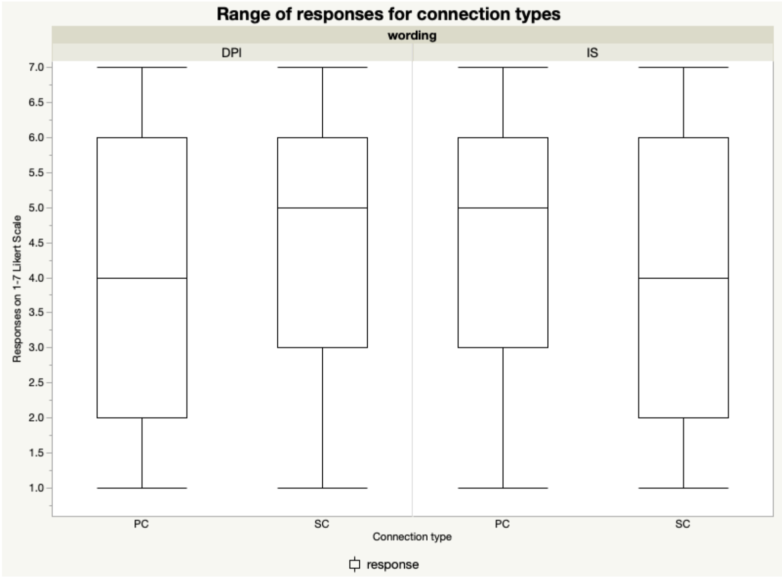


Figure 3.14: Response range of DPl and IS by connection type

3.6.2.5 Results and Discussion

The results (an interaction of wording and connection type) for the French study mirror the results for the corresponding English study (Experiment 1 (DEF-NUM)) with some qualifications. As a brief summary, the asymmetry that was pointed out between the three generic subjects in English was such that, compared to the English BP, IS subjects *increased* expectations that a property was principally connected to the category, but did not affect the judgements about statistically connected properties. Asymmetrically, English DS subjects *decreased* expectations that a property was statistically connected to the kind, but did not affect the judgements about principally connected properties.

The main effects of this study showed that, unlike in almost all English pseudoword studies, the differences in wording were significant (whereas the only main effects of wording in the English pseudowords studies were in Experiment 3 (DEF-NUM) and Experiment 5 (DEF-NUM)). Conversely, there was no main effect of connection type, while all English pseudoword studies displayed a main effect of connection type.

The results from the French version are comparable to the studies in English to an extent, in that properties that were trained with IS subjects are more likely to be expected to be principally connected to the category than those trained with DPl subjects. This was only the case for some of the English pseudoword studies. Unlike in the English studies, the IS in French also seems to decrease expectations that expectations are statistical. This is the pattern that was predicted but not observed for the comparison between English BP and DS subjects in Experiment 2 (DEF-NUM), but not for the English IS subjects.

Based on this comparison, the data suggest that there is a stronger asymmetry between DPl and IS subjects in French. This could suggest that the French IS performs largely similar but not identical functions compared to its supposed English counterpart. Given that the English pseudowords studies did not reliably show these interactions, more follow-up studies, similar to those run in English, should be carried out to assess whether this effect is reliable in French. In a sense, the French IS seems to perform both functions, by increasing expectations of PCs and decreasing expectations of SCs at the same time. This, however, is based on the assumption that the French DPl is comparable in its function to the English BP, and moreover, that they are both equally acceptable baselines for comparisons with other generic subjects in their respective languages. While these studies have not provided reliable support for the hypothesis that subtle morphosyntactic differences in generic statements can affect the conceptualisation of the properties predicated of the kind, they have still provided some mixed results using novel methodology. Both the English and French version could benefit from an independent baseline. I will address the results and broader implications of the results of the studies of this chapter in more detail in Section 3.7 below.

3.7 General discussion

The results of these three sets of learning studies do not provide consistent support for the hypothesis that the morphosyntactic form of a generic sentence affects the way individuals encode the connection between a kind and a property and subsequently generalise this connection to further uses of that property under certain circumstances.

- While some experiments show interactions between connection type and wording, others do not.

- Experiment 3 (DEF-NUM), comparing all three generic subjects without a separate training phase, in fact shows a preference for BP subjects with questions targeting PCs, which is the opposite pattern as predicted.

- This raises the question as to how to explain that some studies might show the predicted effects. Based on the findings in this chapter, this could be modulated by the amount of input: Suppose that with careful manipulation and follow-up studies, these interactions are consistently found. This could then suggest that there is evidence that the grammatical form of generic sentences allows us to adjust our conceptual understanding and expectations of these properties accordingly.

- In short, this could provide support for the argument that too little training does not allow participants to become sufficiently familiar with the property and makes it impossible to access the participant's linguistic judgments and intuitions sufficiently. Therefore, these studies show a preference for BP subjects.

- Too much training seems to subject participants to a cognitive overload where they are unable to carry out the task with their natural intuitions anymore. Crucially though, comparing the results from the studies that provided either no training or extensive training, the more distinct results, which are more closely in line with the hypothesised behavioural responses, are those from the extensive training condition.

- Thus, while the main caveat comes from the observation that the reported effects are highly sensitive to the right amount of training, the lack of interaction of wording with connection type within questions targeting SCs in the extensive training condition could be attributed to domain-general limitations.

- More importantly, the findings indicate that while generic sentences might not be exclusively interpreted by world knowledge, the importance of the contribution of the morphosyntactic form on our conceptual system could be overstated. However, these studies are unusual in that there are many pseudowords, which participants might be unfamiliar with as a task type.

- Further, there are no auditory cues that might convey more specific prosodic cues that disambiguate possibly ambiguous readings between generic and episodic interpretations (although this will be more specifically addressed in the next chapter). Specifically, these experiments did not provide consistent support for the theoretical claim that NUMBER directs our conceptual understanding based on the semantic objects the sentence's subject expresses.[11]

One main question that these mixed results leave open is whether we can find experimental support for the theory that languages refer to kinds via two separate mechanisms: the first is via direct kind reference with definite generics, i.e. the numberless DS in English. The second is mediated by NUMBER which instantiates entities, i.e. the IS and BP in English.[12]

The semantics of different grammatical subject types influenced the way participants encoded and generalised the new properties they were learning. In the experiments presented in this chapter, BP subjects acted as a baseline since they are licensed in all generic contexts and they do, in fact, behave similarly across different conceptual conditions. However, unlike the initial predictions, participants judged even BP subjects to be better suited to express principally connected properties, compared to statistically connected ones, despite their hypothesised

[11]However, the cases in which there were interactions between connection type and wording are, to an extent, in line with Prasada & Dillingham, 2006, 2009, though the results for DS subjects were not specifically addressed or predicted in their papers. Prasada suggests that we may characterise kinds via characteristics of a sample of the kind (BP in English), an arbitrary instance of the kind (IS in English) or via exemplification (DS in English) and therefore have distinct morphosyntactic properties of generic subjects corresponding to the underlying conceptual mechanisms (p.c.).

[12]This also still leaves room for the covert modal quantifier GEN, which may act to differentiate sentences with IS subjects with either episodic or generic reference.

ability to refer to both types of properties equally. This is evidenced by a consistent main effect of PCs over SCs.

The experiments presented in this chapter have contrasted BP generics with the IS and DS in English, as well as the IS and DPl generic subjects in French. However, future studies should include a baseline condition other than the BP, precisely because the seemingly dual nature of the BP needs to be explored in more detail. This baseline would ideally be provided by specific reference, such as *this* (e.g. Hollander et al., 2009; Gelman et al., 2010). Additionally, I have referred to the significance that the type of category plays. While the experiments presented here avoid reliance on world knowledge and reduce expectations as to the category type, the morphosyntactic variation exploited in these experiments could be applied to other studies that employ only one type of category, e.g. animal kinds. Based on the high essentialism that animal categories display, the results for DS subjects with principled properties are expected to be higher than in this set of studies, where no such expectations as to the level of essentialism could be made (e.g. Fuellenbach & Gelman, 2019; Fuellenbach et al., 2019). This will be further explored in subsequent chapters of this thesis, particularly Chapters 4 and 5.

Based on these theoretical and experimental findings, future research needs to investigate other properties of generic phrases, such as the role of the predicate in the VP. Our experiments have contrasted BP generics with either one of the numbered generics found in English. Future studies will include a baseline condition other than the BP, precisely because the seemingly dual nature of the BP needs to be explored in more detail. This baseline may either be provided by quantifiers, such as *most* (as in e.g. Leslie et al., 2011), or specific reference, such as *this* (e.g. Hollander et al., 2009).

In light of the versatility of morphosyntactic subjects, the components that license a generic interpretation, as opposed to an episodic one, remain of interest. BP, IS, and DS subjects (with the same surface structure as DN) can all occur in sentences where a generic reading is not available, e.g. *Dogs are/A dog/The dog is in the yard.* Thus, the morphosyntactic cues on the surface do not themselves determine whether the sentence receives a generic reading. Rather, interlocutors must pay attention to a variety of cues distributed over the sentence (if not at discourse level) when determining its interpretation. Similarly, quantified subjects appear to have similar interpretations to generic sentences. Compare the quantified *Every dog is/Most dogs are four-legged* with the generic *Dogs are four-legged*. The former type is often misremembered as a generic expression (e.g. Leslie et al., 2011; Khemlani et al., 2012; Gelman et al., 2010), suggesting that sentences may not exclusively be interpreted in lockstep with their morphosyntactic properties. The way such interpretations arise and the role of non-grammatical factors will be crucial for our understanding of acquisition and use of generic knowledge (Leslie, 2008; Lazaridou-Chatzigoga & Stockall, 2013).

Finally, the underlying theoretical assumptions are cross-linguistically informed and embedded within an emerging body of research on the cross-linguistic differences in the expression of genericity (e.g. Mari et al., 2013, for a comprehensive overview). Languages such as German or Italian, for instance, have a less restricted use of determiners compared to English, while French requires determiners obligatorily in generic NPs. Such observations suggest that speakers of different languages deploy techniques that vary from one another with regard to the overt syntactic and semantic cues employed to express the various types of generalisations. Additionally, cross-linguistic variation may make the acquisition of different kinds of generic phrases more or less straightforward. Future research will therefore further consider other languages and compare the presumed underlying structures, thus further uncovering the functions of NUMBER and DEFINITENESS.

Overall, the nature of the relationship between language and the human conceptual system remains to be explored. The research presented in this chapter suggests that it may be guided and constrained by grammatical devices in language. Aligning grammatical distinctions with their corresponding conceptual differences can provide us with a better grasp of how we come to understand the world when our experience of it is limited compared to the vast amount of our seemingly effortless generalisations.

The tasks carried out by participants in the set of studies in this chapter excluded world knowledge to a large extent. However, a lot of learning makes use of contextual information and is typically not carried out in isolation.[13] Seeing that these experiments have shown an effect of morphosyntactic variation, it raises the question as to how that comes into play in other learning paradigms. Developmental psychology has investigated the role of generic language in studies that use pictures of novel animals with child participants (see e.g. Rhodes et al., 2012; Marchak & Gelman, 2018).

To further investigate the role of world knowledge, and specifically pre-existing (superordinate) category knowledge, I will specifically focus on including this type of knowledge in a match-to-sample task, conveying visual information. My goal for the subsequent chapters is to bring together the findings from linguistic and psychology experiments and see whether the effects that were observed in the studies presented in this chapter, which exclusively relied on linguistic cues, can shed light on how studies that use visual stimuli tap into the conceptual system. I will further test the role of morphosyntactic cues but do so in a more naturalistic setting. A series of studies will combine the effect of varied morphosyntax with other cues, primarily visual ones. This will also allow us to better understand how various types of cues interact in concept acquisition. Finally, these studies will include picture books and aim at more closely imitating a developmental trajectory to follow up some of the initial observations on the role of learning and exposure.

[13]Note that the set-up of these studies did not investigate whether participants in fact did not make any assumptions as to the type of category of the novel "things" that they learned about. While the language in the instructions was deliberately chosen to be as neutral as possible, it is not impossible that participants largely assumed that the novel "things" were, e.g. animals, rocks, plants, etc. However, I argue that none of these distinct categories were formed, based on comparing the results of this set of learning studies to those presented in Chapter 5.

How much wood would a woodchuck chuck
if a woodchuck could chuck wood?

— Robert Hobart Davis *The Woodchuck Song*

4

Reading and hearing about novel animal kinds

Contents

The previous chapter has investigated the role of morphosyntactic cues for our expectations of category-property links. The experiments were conducted to look at the effect of varying NUMBER and DEFINITENESS, in the absence of world knowledge. While the studies consistently showed a main effect of connection type, with principally connected properties receiving higher ratings than statistically connected ones in all of the experiments run in Chapter 3, and crucially regardless of the subject form, the studies were inconsistent in three ways:

- Only three of the eight experiments showed an interaction of wording and connection type (Experiment 1 (DEF-NUM), Experiment 3 (DEF-NUM), and Experiment 4 (DEF-NUM)).
- The remaining five experiments did not show an interaction of wording and connection type.
- The direction of the interactions that were found were inconsistent, with Experiment 3 (DEF-NUM) showing an interaction in the opposite direction as predicted.

The interaction I predicted was that IS and DS generic subjects would favour PCs by rating them higher than SCs, while BP subjects would equally support PCs and SCs. However, the results were not clear and can therefore provided only limited support for the hypothesis developed based on the literature review in Chapter 2. Given that the studies found some interactions, it is likely that the design did not target the morphosyntactic differences effectively. Notably, these studies only provided morphosyntactic information but no world knowledge. It is possible to assume that the lack of superordinate category knowledge or other types of visual information led participants to not carry out the task as predicted.

This chapter will address the effect that the gap of having no world knowledge might have had. The studies presented here will provide a variety of morphosyntactic cues but additionally also provide world knowledge. I will continue to investigate the role of morphosyntactic variation in generic subjects, and how changes in generic subjects affect our expectations of the conceptual make-up of the category. To address this question of methodology, I will present a series of studies that introduce novel animal kinds, accompanied by pictures. In other words, these studies include more information than the *pseudoword* studies. The goal of these studies is to investigate the interaction of multiple types of cues and how morphosyntax influences the conceptualisation of principally and statistically connected properties of animal kinds, eventually including data from child participants to provide a developmental perspective.

4.1 Match-to-sample tasks: wording, properties, and cues

How do we generalise properties of novel animal kinds when we only have very limited cues available? One way of investigating this question is to use match-to-sample tasks. Also known as novel word extension tasks, match-to-sample tasks have been used to understand how adults and children generalise properties based on the exposure to one instance. The underlying assumption of this task is that we normally group members of a kind based on their overall visual similarity. Hollander (2007, p. 54), who uses this paradigm in a study investigating the role of generic language, refers to Baldwin (1992) and Imai et al. (1994) who provide independent evidence for a shape-bias in categorisation tasks and novel word learning. The reasons for *shape* acting as something akin to a default categorisation mechanism are not fully understood yet. However, ultimately they must be tightly linked to the mechanisms of acquiring world knowledge via statistical exposure, forming prototypes, or via access to mini-theories that are built over time. In a somewhat related series of experiments, Smith (1989) tested the assumed *shape* preference through asking participants to group triples of circles, where two circles were overall more similar, while one of those and the third one were identical on one dimension but different on another. To illustrate, two circles might share the same colour but differ in size, or vice versa. Smith (1989) found developmental differences, such that children tended to group circles sharing overall similarity, i.e. somewhat similar in colour and size, but not identical on either of these dimensions, while adults grouped the circles that were identical on one of the dimensions. Seeing that a *shape* match, or Smith's 'overall similarity match', was taken as the baseline, it will be important to keep in mind children's developmental trajectory for the interpretation of the results.

A thorough review of which of these theories seems to be the forerunner to explain kind acquisition will be given in Section 6.2. However, note that the shape of an animal kind could easily be hypothesised to be one of the kind's principally connected properties. At this point, it suffices to rely on the assumptions made in earlier studies, which have been carried out to investigate whether these results are replicable. With many studies demonstrating that visual similarity can be a sufficient cue to form a novel category, I rely on the assumption that this task type provides at least some insight into generalisation processes.

4.1.1 The role of labelling

The ability to make inferences based on visual input seems to be language-independent. This does not mean that the process of conceptualisation, and establishing the nature of category-property link, does not highly emphasise linguistic cues. Rather, one might draw a parallel to the role of (caregivers') recasts in first language acquisition: They are a useful source of input for children and aid their linguistic development, yet they are not necessary for a child to acquire a full grammar (O'Grady, 2005, p. 169). Therefore, taking into consideration the powerful effect of labelling itself, this provides a useful testing ground for the hypothesis that generic language can bias people's generalisations. This has also already been observed by Hollander (2007), referring to a wide range of studies such as Gelman & Coley (1990), specifically addressing the role of labeling, but also Kemler Nelson et al. (2000), Booth & Waxman (2006), and Diesendruck et al. (2003), demonstrating how people may overcome shape biases in artefacts if two objects share a function.

Booth et al. (2005)'s study focused on the developmental trajectory of concept acquisition and highlights the role of word learning therein. They used a match-to-sample paradigm which included a target object, an exact match, three objects which differed from the target in texture, shape or size, and a distractor object. Their experiments showed that children upwards of four years of age, extend categories based on the labels provided. Crucially, they found differences in participants' willingness to extend category membership to a novel object if it differed in texture, based on whether the novel category was animate. A difference in texture between the target and the novel object lead to lower acceptance rates if the category was animate. Booth et al.'s experiments provide support for the hypothesis that word learning and category knowledge interaction in concept acquisition.

Therefore, the specific (generic) wording in a match-to-sample paradigm may be a powerful tool to steer people away from their initial bias in favour of visual similarity and towards the characteristic that is explicitly mentioned with regards to a novel kind. Systematic variation in the wording of the generic subject can be crucial in establishing precisely what is the role of generic subjects. In other words, if people are forced to choose between two options, they might choose an instance that possesses this predicated property based on the wording with which it is introduced. In these cases, they might prefer it over the visually more similar instance.

This overview demonstrates that match-to-sample tasks typically bring together two types of stimuli that enable acquisition of category knowledge: images and words. When people are tasked with choosing novel members of a kind, they often need to engage with competing input. Without labels, the most powerful cues are the visual descriptors of a new kind, based on which novel category members can be identified. This can change when labels are given to new kinds and categorisation becomes a combination of paying attention to both the image and the accompanying linguistic cues.

In this context, it is important to acknowledge that visual stimuli are inherently connected to the wider mechanisms of the cognitive system. Specifically, they evoke similarities to other kinds about which we already possess categorical knowledge. In most match-to-sample tasks, the novel kinds that are introduced are animals. Without knowing anything about these new animals, participants at least know something about animals in general. Moreover, they likely have specialised knowledge about mammals, birds, reptiles, etc. To fully disentangle pre-existing knowledge, which could further be based on linguistic knowledge, we need to assess these pieces individually.

With this in mind, this chapter will look at the effect of comparing the full range of linguistic variables, and will join them with visual stimuli, as used in previous match-to-sample tasks. It will pay special attention to how the evocation of images of animals affects the accessibility of distinct generic interpretations, which linguists have referred to based on minimal pairs. These were the basis of the experiments in 3. Section 4.2 lays out the basic design for the six experiments presented in this chapter. It will do so by explaining the assumptions and findings in Hollander et al. (2009)'s *kevtas*-studies. Sections 4.2.1-4.2.3 present the results of the first three studies, which tested adult participants. Recall that both experimental settings as well as naturalistic settings, i.e. those with spontaneous utterances, often include a subtype of linguistic input: spoken, as opposed to written, words. As an experimental variable, they need to be carefully manipulated. I will discuss the design of audio-recorded and edited generic phrases in Section 4.3. Finally, Section 4.4 takes the findings from Experiment 7 (world knowledge)-Experiment 9 (world knowledge) and translates the match-to-sample tasks into a paradigm that allows for child participants. Based on the results of these three studies, I discuss how developmental patterns vary for the three generic subjects in English. Section 4.5 summarises the findings from the six studies in this chapter. It also addresses how we can take our new insights and transpose them into an experimental paradigm that more naturally reflects the acquisition of world knowledge for children.

4.2 Do you know about kevtas?

The overall aim of this chapter is to bring more closely together the experimental methodology of psychology studies, especially those looking at the development of generic interpretations, with linguistic assumptions about genericity. This can provide us with a unique perspective on how the form of a linguistic subject affects how we conceptualise the connection between a kind and its properties. Moreover, by looking at the developmental aspect, we can make a more educated guess as to whether any or all interpretations of genericity are innate or learnt. I thereby address the debate between *Generics-as-Default* and GEN from a new angle. Specifically, I will be looking at the results of more fine-grained stimuli in the context of well-established acquisition studies, much in line with the design in Chapter 3.

To start with a broad overview, the studies in this chapter explore assumptions about the accessibility of various generic interpretations. Ultimately, to understand whether there are default generalisations, and if so, if all default generics are created equal, we need to compare the available generic subject forms with one another. This chapter will make use of stimuli that have previously been used and have shown that we use generic language differently from specific language. Both children and adults were found to be sensitive to the effects of generic subjects, albeit the only type tested was the bare plural (BP). To pick up on the idea of learnability explored in Chapter 3, the experiments here were initially run on adults, with the goal of testing child participants as well. Including children of a broad age range enables us to draw developmental comparisons. Such evidence will help us further constrain theories of development and learnability, and underpin the findings from the corpus analysis and addresses the differences in theoretical assumptions, specifically with regard to *Generics-as-Default*.

Match-to-sample tasks provide a simple yet powerful experimental paradigm that targets people's understanding of novel categories. Hollander (2007, p. 55) used this paradigm to investigate whether, and if so how, children employ generic language to focus on the significance of a property to its category. By providing participants with a *target* image of the novel kind, which is introduced with a generic label, in this case a BP subject, a property of the novel category was highlighted. Building on the match-to-sample paradigm which uses a triplet of images, participants might be told that a novel animal kind named *bant* 'is striped'. They then have to answer the question, "Which one of these is also a bant?", choosing between an image that is a *shape*-match, i.e. overall displays high shape similarity, and a *property*-match, which is also striped but is varied perceptually from the target image on at least two salient features, one of which was the colour of the animal. Hollander ensured that the *property*-match was still recognisable as a potential member of the kind by keeping them similar enough (based on the findings in Davidson & Gelman, 1990).

This question is closely aligned with the research question of this thesis regarding the (behavioural) effects of the distinct semantics of generic subjects. There are two crucial differences between the studies presented in this chapter and those run by Hollander et al. (2009): First, they did not investigate generic language beyond the BP. Second, their research addressed whether properties are highlighted as essential to a category through generic language. By contrast, I focus on the type of connection between category and property, already working under the assumption that they are essential to it if introduced with a generic subject.

To follow up on Hollander et al's studies with these more specific questions, it was crucial to make an informed decision about this experiment's stimuli early on in the design of the studies. The overall aim in this process was to join the take-away points from Experiment 1 (DEF-NUM), which disentangled the effects of principled vs statistical properties (PCs and SCs, respectively) with the picture-book design from well-established studies in developmental psychology. Initially, I consulted the stimuli in Prasada & Dillingham (2006, 2009) to determine which of their previously used properties were appropriate for PCs and SCs in this adaptation of Hollander et al. (2009). In those studies, participants rated generic statements with normative explanations and *by virtue of*-paraphrases. PCs should lead us to accept these explanations and paraphrases and judge them as more natural sounding (as in (1)) than generic statements based merely on SCs (as in (2), both from Prasada & Dillingham, 2006):

(1) a. Dogs are four-legged.
    b. Dogs, by virtue of being the kinds of things they are, are four-legged.
    c. Dogs should be four-legged.

(2) a. Taxis are yellow.
    b. ?Taxis, by virtue of being the kinds of things they are, are yellow.
    c. ?Taxis should be yellow.

When looking through the ratings for SCs, there are much higher acceptance ratings for the *by virtue of* and *is one aspect of* paraphrases whenever the stimuli included animal kinds compared to those given for artefacts. This means that the inclusion of animal kinds seems to predict a higher expectation of a connection to be principled. By extension, the inclusion of artefacts seems to favour SCs.[1] A likely explanation for this trend is that category knowledge about animal kinds, in particular knowledge about an animal's diet, seems to strongly imply that the connection be principled. Even when the foodstuff itself is not connected in a principled manner to the animal kind in question, the overall mechanism seems to override this more specific datapoint, which is merely statistically connected. Taking advantage of this finding, and trying to identify strong predictors of these two types of connections under investigation, the stimuli were designed so that connections that were intended to be understood to only be statistically prevalent included artefacts. To wit, instead of using *Cats drink milk* or *Winters are snowy* which Prasada & Dillingham (2006) argue to be statistical properties of animal kinds, the stimuli for the experiments presented in this chapter use properties such as *throw glow sticks* or *give rides at the petting zoo* as SCs.

Thus, in the development of properties that are likely interpreted to be principally connected, three factors were considered: first, properties were chosen based on those that already scored high in the ratings provided in Prasada & Dillingham (2006). Second, they needed to be easily adaptable to be able to extend the stimuli used in Hollander et al. (2009). This was easily done as Prasada & Dillingham had an extensive list of stimuli from which only a subset of eight was necessary for the *kevtas* studies. Third, as explained above, only animal categories were included for these studies, and therefore artefact and social kind categories from these previous studies were disregarded in the selection.

The following sections introduce experiments in which I extended the assumptions and findings from Hollander et al. (2009), and created studies that take the

[1]As an analogy provided by real world examples, we find that the same property of 'being red' evokes a principled connection when combined with the natural kind 'robin' but a statistical connection with the artefact 'barn':

(i) Robins are red.

(ii) Barns are red.

existing paradigm of comparing BP generic subjects with specific wording (with the demonstrative determiner *this*) further. I developed a stimuli set that accounts for the three-way distinction of generic subjects of count nouns in English. Following up on Hollander et al. (2009)'s previously tested design, I hope to target the effect of morphosyntactic variation in the subject DP on the conceptual system from a slightly different angle. The results give us more insight into generic knowledge and how it is structured and accessed in comprehension studies.

4.2.1 Experiment 7 (world knowledge): Adult participants

This experimental series was designed to further test the hypothesis that the morphosyntax of the subject can raise expectations as to whether the connection between concept and property is principled or statistical in a targeted manner. The design of the study was adapted from Hollander et al. (2009), whose assumptions and findings were presented in Section 4.2 above. I used a between-subjects design, such that every participant is assigned one wording condition and one connection type. Consequently, no participant is exposed to more than one type of subject form. More specifically, participants saw a picture of a novel animal creature and were provided with a property that this creature possesses (e.g. *Kevtas have woolly fur* or *A febbit wears scarves*).

After seeing this picture of a novel animal, participants were asked to identify another animal of the same category from two pictures. Based on Hollander et al. (2009)'s studies, the design of the two answer choices is such that one is overall more similar in *shape* but does not possess the predicated property, while the other choice is overall more dissimilar but displays the *property* the target critter has. As pointed out above, referring to Hollander (2007, p. 70), this is a stringent test of the wording effect, as "children are particularly attentive to shape in their early categorizations and word extensions" (with reference to Baldwin, 1992; Smith et al., 1996; Landau et al., 1988).[2]

[2]However, this notion will be challenged in Chapter 6, in which I point out the differences between comprehension and production-like studies as well as questioning the validity of a *shape* match as a baseline against which to compare *property* matches. To foreshadow these findings, I

I predict that variation in generic language will show differential results, such that the BP (*kevtas*) leads to high generalisability and stability scores with so-called principled properties about the animal (e.g. *having red wings*) as well as merely statistically connected properties (e.g. *throwing glow sticks*). In line with previous hypotheses, the IS (*A kevta*) should lead to high generalisability with principally but not statistically connected properties. The DS (*The kevta*) is predicted to lead to lower generalisability with statistically connected property, but not necessarily to higher generalisability with principled properties. This is similar to the IS, but expected to render lower scores overall, due to the more restricted use of the DS. All these predictions are following the theoretical assumptions made for the experiments in Chapter 3. By contrast, specific language (*This zarpie*) is not expected to lead to a high occurrence of generalisations of properties as it conveys specific and not general information about the new animal. It is used as a baseline against which the three generic subjects are compared.

4.2.1.1 Participants

426 adult native speakers of American English were recruited via Amazon's MTurk. They were randomly assigned one of 8 conditions (n $\approx$ 53 per condition). Participants provided informed consent and received financial compensation ($0.50) for their participation, which took approximately 5-10 minutes.

4.2.1.2 Items

The stimuli for this study were composed of those stimuli from Hollander et al. (2009) that were suitable to be adapted for the slightly different purpose of this study. Other stimuli were adapted from previous studies using novel animal creatures in the Conceptual Development Lab, as well as some additional stimuli that were designed specifically for this study to include a wider range of properties. This concerns the inclusion of artefacts to represent statistically connected properties,

argue that *shape* can easily be interpreted as a principled property of an animal kind and can be difficult, but not impossible, to separate conceptually from another, overtly stated principled property. Of course, given the well-established nature of these studies, *shape* and *property* cannot be equated and seem to still perform distinct functions.

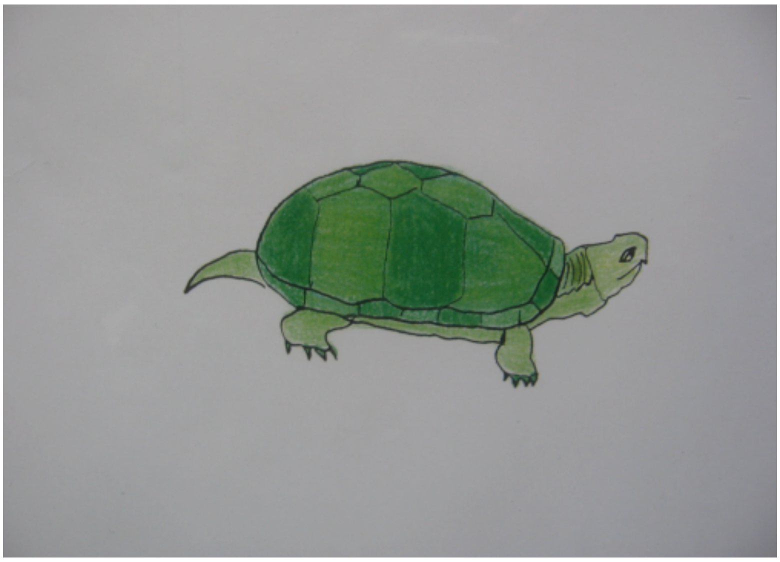


Figure 4.1: Sample target stimulus, in PC-BP condition. Introduced with the prompt: "Do you know about plogs? Plogs have green shells."

which, as explained above, are a novel consideration. When designing new stimuli, the original procedure of creating pseudoword names of two to three syllables in length, counterbalanced across groups (defined here by connection type) was followed. A total of 16 triplets was designed, half of which were used for the PC condition, and the other half for the SC condition. Each set consisted of a triplet of pictures as illustrated in Figures 4.1-4.3. To illustrate, the target picture for *plogs* was introduced with the prompt *Do you know about plogs?*, and a sentence introducing its property, which was either principally connected, as in *Plogs have green shells*, see Figure 4.1, or statistically connected, e.g. *Kevtas wear scarves.*

The target picture was followed by two more pictures, one in the *shape* condition (Figure 4.2) and the other one in the *property* condition (Figure 4.3). The former is overall more similar in shape to the target but lacks its predicated property, while the latter is overall more dissimilar in shape but displays the specific property of the target stimulus. This resulted in a 2 (connection type: principled vs statistical connection) × 4 (wording: BP vs IS vs DS vs 'this') between-subject design.

Some of the properties included (multiple) countable parts, e.g. *have green shells* or *wear scarves.* These could theoretically either be singular or plural for singular-numbered subjects. These properties occurred in the plural, rendering *A plog wears*

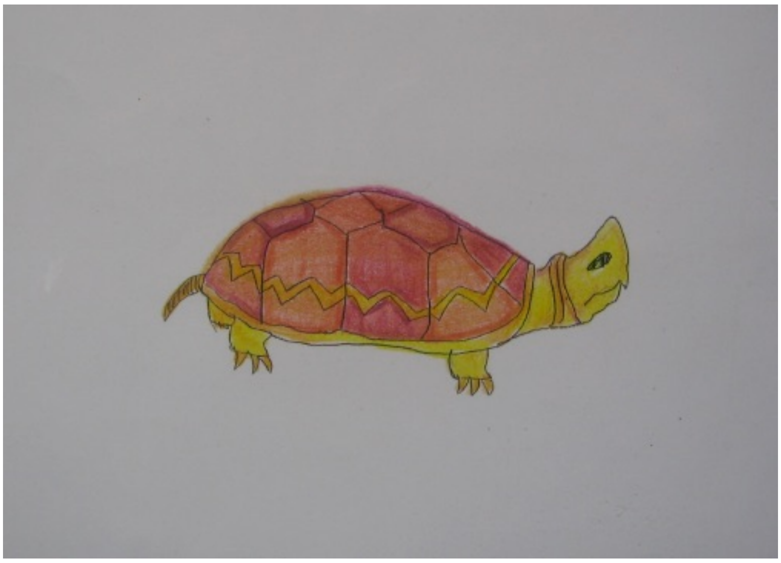


Figure 4.2: Sample shape stimulus, provided as one option for the question: “Which of these is also a plog?”

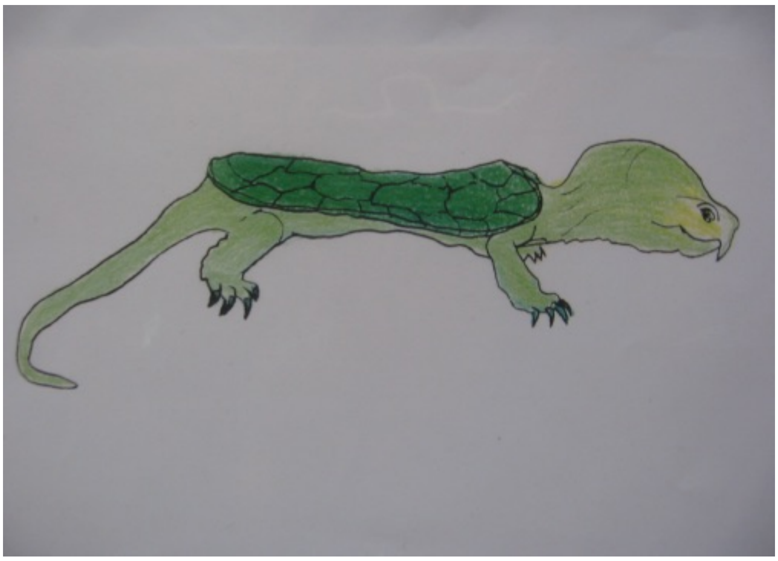


Figure 4.3: Sample property stimulus, the other option for the question: “Which of these is also a plog?”

scarves instead of *A plog wears a scarf*. The reasoning for this is two-fold: First, it should make a generic interpretation more accessible to participants. Second, the property in its plural form renders the overall stimuli more similar across conditions. Compare *A plog has green shells* to *A plog has a green shell*, where the latter property sounds like a predication of detachable parts or a possessive relationship, rather than a descriptive phrase of the animal's physical constitution. See Appendix C.5 for a full set of the stimulus pictures and sentences in the BP condition and Appendix C.5.1 for sample stimuli of the IS, DS, and control condition.

4.2.1.3 Procedure

Participants were recruited via Amazon's Mechanical Turk (`https://www.mturk.com`) and were redirected to the study which was hosted on Qualtrics (`https://www.qualtrics.com`). First, they provided informed consent to participate and filled out a short linguistic questionnaire (see Appendix C.3 for full questionnaire). Participants were randomly assigned one of eight conditions.

At the start of the experiment, participants were introduced to the task with the following information:

> Welcome!
>
> You will be seeing pictures of some new animals. These are different from the animals we know, so we have made up new names for them.
>
> For each one of the new animals, we will also ask you if you can identify another one. We will show you two pictures to choose from and you can just click on one of them.

During the experiment, they were asked to click on either one of the two image choices presented below the target animal. They were each presented with 8 trials, consisting of the target-shape-property triplets, presented in random order. One was the choice similar in *shape*, the other looked dissimilar but possessed the predicated *property*. The order in which the two answer choices were presented was also randomised. All of these randomisations were evenly presented across participants.

After the eight test trials, participants received a break message explaining that they were almost done and were then presented with two memory tasks, showing two random target pictures and asking them to write down in full sentences what they learned about this animal.

4.2.1.4 Analysis and coding

The data of 426 participants were collected. Initially, the data of all the participants were analysed. However, due to a surge in questionnaires being filled out by bots and/or low-quality workers during the period of data collection, multiple approaches were taken to establish whether the collected data was collected from human participants, whose native language is American English. Common indicators of bots responding to online surveys are all-capitalised comments (especially one-word comments, such as “GOOD”), but these are by no means a guarantee that these data stem from bots or, equally, all data entered by bots are excluded.[3] Normally, it is possible to look for repeated IP addresses or GPS-locations. However, due to data protection laws in Europe (GDPR 2018), I was unable to collect such information, since data collection took place in the U.S. but was analysed and stored in Europe. Bearing these precautions in mind, I excluded the data of 61 participants based on, e.g. either having very fast response times, only having entered all-capitalised and/or one-word comments, or having provided the same responses for all trials. Therefore, the analyses presented below are based on the data of the remaining 365 participants.

Answers for the *shape* option were coded as “0”, *property* answers were coded as “1”, in other words, the data for the analysis are binary variables. Participants choosing the picture with the *shape* match is interpreted as participants choosing the baseline option, which participants prefer in the absence of any generic of non-generic information about the triplets of images. Participants choosing the picture with the predicated *property* is interpreted as participants generalising the information in the prompt, and consequently that their generic interpretation overrode the assumed *shape*-bias. Every participant saw eight trials, and thus they could get a minimum score of 0 and a maximum score of 8. The scores are averaged for each condition. This type of scoring and analysing was chosen in

[3]Note that the comments from this study were primarily collected to check whether participants found the task intuitive and were taken into consideration for the design of subsequent studies.

line with the original analysis of the match-to-sample studies in Hollander (2007); Hollander et al. (2009). However, going forward, this methodology should be revised as the input into the model would benefit from being a proportion or percentage score instead, rather than the raw scores of the binary outcomes from choosing one picture or another. In the case of this thesis, however, all analyses follow the same methodology as in the original papers, using binary variables based on the assigned scores of “0” for *shape* and “1” for *property*.

A 4 (wording: BP vs IS vs DS vs ‘this’, between participants) × 2 (connection type: PC vs SC, between participants) ANOVA was conducted over scores of *property* > *shape* response options to establish whether the differences between the scores for generic subject types for each of the property types were significant. This analysis revealed a main effect of wording ($F(3,418) = 8.527$, $p < .0001$) and a main effect of connection type ($F(1,418) = 56.197$, $p < .0001$). The analysis did not find an interaction between wording and connection type ($F(3,418) = 1.308$, $p = .271$). These effects are visualised in Figure 4.4.

Due to a lack of interaction in the omnibus model, no paired comparisons were carried out.

4.2.1.5 Results and Discussion

The purpose of this experiment was to investigate a possible interaction between the morphological form of a generic subject and the kind of property it expresses. While the results show main effects of wording and connection type, they do not show an interaction. Thus, the data do not show that the specific morphological form of a generic subject affects our interpretation of a property connected to a novel kind, but they show that all three generic subjects scored significantly higher than the non-generic condition.

For PCs expressed with the specific label ‘this’, their scores are similar to those of generic subjects associated with statistically connected properties. I argue that the main effect of PCs in generic subjects is due to the higher generalisability of those properties to the entire kind, compared to that of merely statistically

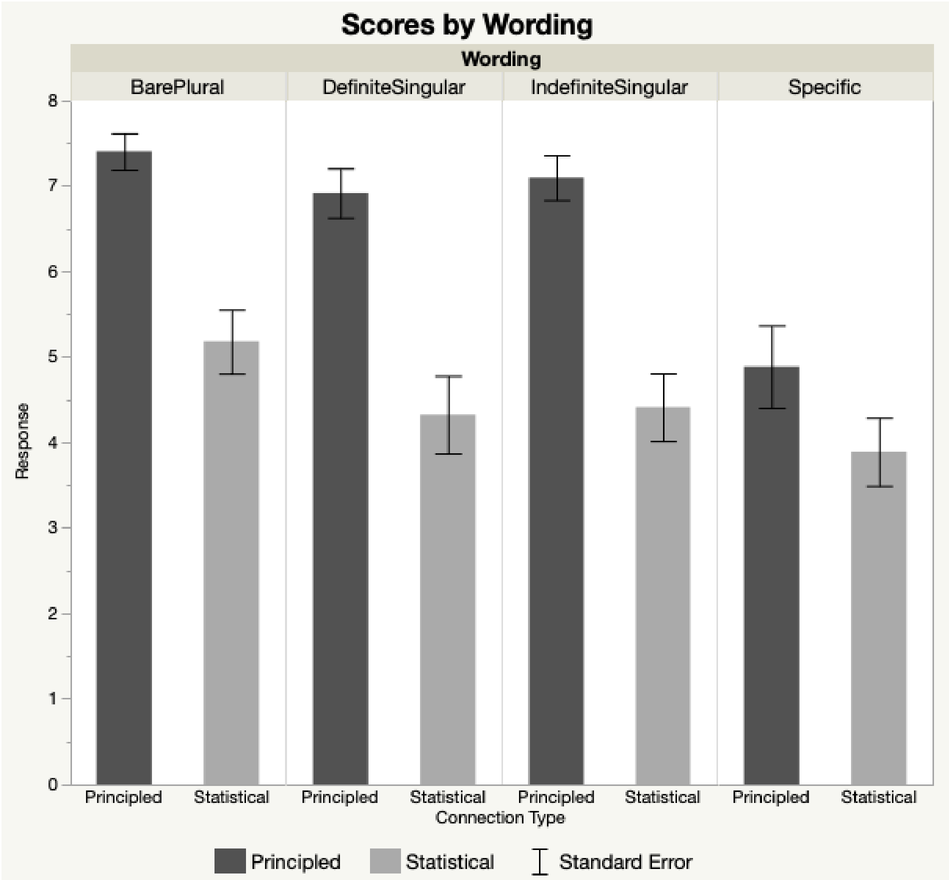


Figure 4.4: Averages of choosing property over shape matches for predicated properties by wording. Error bars = ± 1 SE from the mean. Principled properties indicate, e.g. *have green shells*, statistical properties indicate, e.g. *wear scarves*.

connected properties. These statistically connected properties that are expressed with a generic subject are perceived as essential to the kind as properties that are principally connected but are not expressed by a generic subject, i.e. are only known to be true of one specific instance.

In particular the high scores of PCs for all generic subjects are in line with the hypothesis and findings in Hollander et al. (2009), who argued that richly structured animal categories have high essentialism when properties are predicated with generic language. These findings were extended to IS and DS subjects. It then follows that these studies can demonstrate that, even with minimal input, the DS subject can be interpreted generically, just like the BP and the IS. To the best of my knowledge, this was the first learning study of this type that systematically pitted BP, IS, and DS subjects against a specific-label with a match-to-sample task. It is therefore a novel finding that the DS subject behaves just like the other generic subject forms. This study further provides supporting evidence that even with minimal input, DS subjects are interpreted generically, just like the BP and the IS.

Given the differential semantic representations of NUMBER and DEFINITENESS, I had hypothesised an interaction between subject form and connection type. To investigate whether the lack of such an interaction was due to the experimental design, I carried out a follow-up study with some minor changes to the design, specifically addressing the a priori assumptions of the *shape* vs *property* choice, as well as the homogeneity of the stimuli across conditions.

4.2.2 Experiment 8 (world knowledge): Adult participants II (normed)

To ensure that the stimuli were not inherently biased or that there were no large discrepancies between the acceptability of the triplets, a norming study was run before running Experiment 8 (world knowledge) itself as the replication. This norming study examined the individual stimuli by asking participants to choose a match for the target image, thereby providing a more independent measure of the underlying scores of similarity for the image triplets. Participants were introduced to the task by the following message:

> Welcome!
>
> In this task, you will be seeing pictures of some new animals. For each, you will learn the name for that kind of animal. Then you will be asked to select which of two other animals has the same name. Below each animal, you will also see a small picture of something in the context.
>
> Please answer to the best of your ability.
>
> When you're ready, please turn off any distractions, such as your cell phone, for the duration of the experiment! Additionally, please avoid using the mobile version of this site.

During the task, participants were only presented with the specific version of the attributing sentence, e.g. *This is a kevta* and the target image of the novel animal. Participants neither saw the initial prompt (e.g. *Do you know about kevtas?*) nor did they receive information about their properties (e.g. *Kevtas have wooly fur*).

Based on this one sentence alone, they had to choose between the *shape* and the *property* samples, answering the question *Which one of these is also a kevta?*

If we assume the preference for *shape* similarity to hold, then participants should choose the image that we had provided as the *shape*-option as opposed to the novel critter that is dissimilar in shape but possesses the predicated *property.* Recall that Hollander et al. (2009)'s stimuli were designed to vary on multiple dimensions (i.e. colour, texture, colour-pattern, salient and/or disproportionately-sized body parts, and number of parts; Hollander et al., p. 487). This was done while still keeping the *property*-match similar enough that it is a plausible member of the same category. Since the stimuli were amended here, a norming study helps verify that the stimuli are still fit for purpose.

In two norming studies with 50 participants each, 9 triplets were identified that were reliably chosen for a *shape* match (a second norming study was run after the first norming study provided an insufficient number of triplets. Some new stimuli were designed and tested in the second norming study).[4] To have a balanced set of land- and water-dwelling creatures, the following subset of 8 novel animals was chosen: bactras, feps, kevtas, screds, bants, febbits, plogs, and vorzyds. The names were also rearranged such that each set was balanced between mono- and bisyllabic novel names.

Each of these triplets was redesigned such that it was possible to use them in the PC and the SC conditions. To accommodate this change, little images of statistically connected properties were added to all animals. This means that the *shape* match possessed neither the principally nor the statistically connected property, but was an image that was reliably rated as an another instance of the kind in the absence of any further generic or specific information in the norming studies. By contrast, the *property* match looked more dissimilar from the target image, but had both the principally and the statistically connected property.

[4]This means that some stimuli were amended until participants consistently chose the *shape* match. In practice, this simply implies that not all stimuli from Experiment 7 (world knowledge) were included in this study but replaced by updated version. For a full list of the current stimuli, including the images, see Appendix C, Section C.5.

4.2.2.1 Participants

412 adult native speakers of American English were recruited via Amazon's MTurk. None of them had participated in Experiment 7 (world knowledge).

4.2.2.2 Items

In this study, 8 triplets were chosen, based on the results of the two norming studies. They each consisted of a *target* image, a *shape* match, and a *property* match.

By adding little pictures of statistically connected properties, each of these triplets was designed such that it could be used in the principled connections or the statistical connections conditions and thus made the design fully counterbalanced. Participants saw the same stimuli regardless of the condition they were assigned. To illustrate, this means that participants will see a small image of a statistically connected property, e.g. an image of a bottle, even if they are told the animal's property is *being striped*. This was to ensure that there were no systematic differences between participants due to the stimuli that they saw. Thus, there was one *shape* match that possessed neither the principally nor the statistically connected property and one *property* match that looked dissimilar but had both the principally and the statistically connected property.

Participants were assigned a wording condition, but unlike in 4.2.1, they were shown both connections types. This results in a 4 (subject form: BP, IS, DS, control 'this', between subjects) × 2 (connection type: principled, statistical, within subjects) mixed design. The order was random and the order of the set of critters to principled or statistical properties was rotated between the two sets (Set A: bactras, feps, kevtas, screds; Set B: bants, febbits, plogs, vorzyds).

4.2.2.3 Procedure

After consenting to participate in the experiment and confirming that they are a native speaker of English, participants started the experiment. They first read the introduction to the experiment, which was slightly adapted from 4.2.1 (changes highlighted in **bold**):

> Welcome!
>
> In this task, you will be seeing pictures of some new animals. They are different from the animals we know, **so you will learn the name for each of them**. For each animal, **you will see a description and a small picture of something in the context**.
>
> **Then you will be asked to** identify another one. **For this,** we will show you two other pictures to choose from.
>
> Please answer to the best of your ability.
>
> **When you're ready, please turn off any distractions, such as your cell phone, for the duration of the experiment! Please avoid using the mobile version of this site.**

Participants then responded to 8 trials, which consisted of two blocks, although there was no obvious break for the participants. They saw four trials of novel animals with principally connected properties, and four trials with statistically connected properties. The order within the blocks was random, as was the sidedness of the presentation of the two possible answer choices (*property* vs *shape*). The procedure was created to be applicable to child participants as well.

After the testing period, participants were asked to fill out an optional Essentialism Questionnaire, which was a subset of 10 questions from Gelman et al. (2007). The final part was a Demographics section that asked participants to respond to 10 questions about their sociolinguistic background. There was no memory component in this version.

4.2.2.4 Analysis and Coding

The data of 17 participants were excluded as they did not finish the study or had extremely slow response times. The data of the remaining 395 participants were analysed. The data were coded in a similar way to the data in Experiment 7 (world knowledge), with the exception that connection type was now manipulated within participants, and not between. Therefore, participants could receive a minimum score of 0 and a maximum score of 4 for PCs and for SCs, with one point for each trial in which they chose the *property* and not the *shape* match (as opposed to a total score of 8 in Experiment 7 (world knowledge), where connection type was not manipulated within participants. Recall that the data for the analyses were binary, following the original methodology of Hollander (2007) and Hollander et al. (2009), and not percentage scores as should be considered for future studies.

A 4 (wording: BP vs IS vs DS vs 'this', between participants) × 2 (connection type: PC vs SC, within participants) × 2 (block order: PC-first vs SC-first, between participants) mixed ANOVA was run to analyse whether the differences between the scores for generic subject types for each of the connection types were significant, and whether there were any significant differences based on which connection type was presented first. This analysis indicated a main effect of connection type ($F(1,387) = 11.966$, $p < .0001$), a main effect of wording ($F(3,387) = 31.632$, $p < .0001$), but no main effect of block order ($F(1,387) = .155$, $p = .694$). This analysis indicated no significant interaction of wording × connection type ($F(1,387) = 1.070$, $p = .362$), no significant interaction of connection type × block order ($F(1,387) = 1.740$, $p = .188$), a significant interaction of wording × block order ($F(3,387) = 4.301$, $p = .005$), and no significant interaction of wording × connection type × block order ($F(3,387) = 2.305$, $p = .076$). These results are visualised in Figure 4.5.

Due to lack of interactions in the omnibus model, no planned planned comparisons were carried out (apart from the wording × block order interaction, which was not of particular to this research, although an initial exploration indicates that BPs received higher scores when PCs preceded SCs and, conversely, 'this' received higher scores when SCs preceded PCs).

4.2.2.5 Results and Discussion

The results for this experiment (main effect of wording and of connection type, a trend towards an interaction of wording × connection type × block order, but no interaction of wording × connection type or of connection type × block order) replicate the overall findings of Experiment 7 (world knowledge). However, note that the means of the different wording conditions are more distinct in terms of the anticipated interaction, compared to the means of the different wording conditions in Experiment 7 (world knowledge). Specifically, the mean for IS subjects expressing

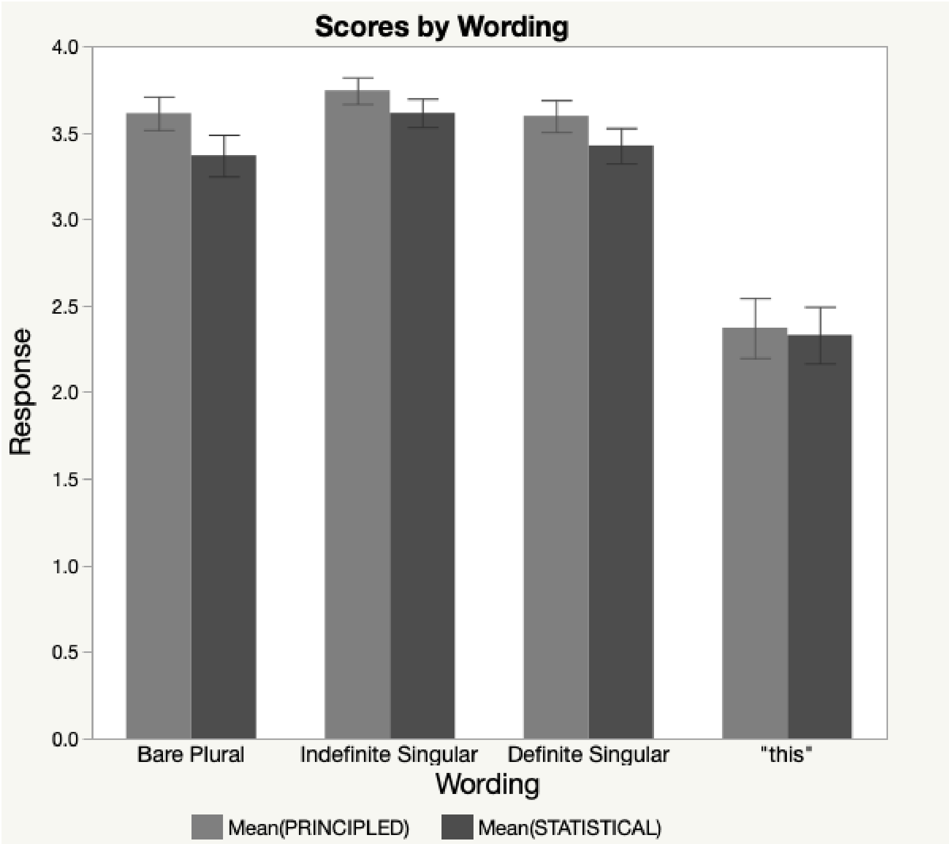


Figure 4.5: Averages of choosing property over shape matches for principled and statistical connections by wording. Error bars = ± 1 SE from the mean

PCs is higher relative to the BP, as well as higher in absolute terms than the IS scores in Experiment 7 (world knowledge) (92% vs 80%).

This is most likely due to the norming of the stimuli, which allows this study to provide a stronger test for the effect of generic language. In the absence of any cues, participants chose the *shape* match as opposed to the *property* match. In the control condition, participants chose the *property* match around 60% of the time, regardless of the connection type that was predicated of the novel animal. By contrast, with generic language, all three generic subjects led participants to choose the *property* match at least 87% of the time. In both *kevtas* studies, the BP and the newly tested singular generic subjects IS and DS behaved similarly. This provides support for the hypothesis that all three generic subjects do not automatically lead to episodic interpretations.

Yet, the interactions for singular generic subjects with PCs and SCs that were predicted were not found. Recall that there were no interactions in Experiments 1-5 (DEF-NUM), with the exception of Experiment 1 (DEF-NUM), Experiment 3 (DEF-NUM), and Experiment 4 (DEF-NUM). One reason for the lack of interaction in the *kevtas* studies might be the language of the initial generic prompt. It asks *Do you know about kevtas?*, which means that the prompt already introduces a novel kind in one of the three generic subjects, namely the BP.

In order to address the inconsistency of exposing participants to two different generic subjects when in the IS or DS condition, I designed Experiment 9 (world knowledge) to investigate whether not using any generic subject in the initial prompt allows for the finer differences between the plural and singular generics to come out in this experimental design. Experiment 9 (world knowledge) therefore tests the same hypothesis with the same stimuli, but crucially without providing the initial prompt that was used in Experiment 7 (world knowledge) and Experiment 8 (world knowledge).

4.2.3 Experiment 9 (world knowledge): Adult participants III (single word label)

Experiment 9 (world knowledge) was designed to specifically investigate the effect of the initial prompt. Both Experiment 7 (world knowledge) and Experiment 8 (world knowledge) used sentences of the form *Do you know about kevtas?* at the start of each trial. This wording was chosen to bias participants towards a generic reading (and to be tested for the use in the child studies to ensure that the children pay attention at the start of the next trial). However, the wording itself might not only facilitate a generic interpretation of the ensuing statement about the novel critter, but also prime participants towards a BP generic reading since the novel kind is first introduced at this point and, crucially, already in one of the three generic subject forms that is used as a variable. Experiment 9 (world knowledge) removed this initial prompt so that this study does not present participants with more than one generic subject form during the studies (recall that wording is manipulated between, not within subjects).

4.2.3.1 Participants

410 native speakers of American English were recruited for this experiment. None of them had participated in either of the previous experiments.

4.2.3.2 Items

The stimuli are the same as in Experiment 8 (world knowledge), with one change: None of the generic statements are preceded by the initial prompt *Do you know about...?*. Thus, participants saw prompts such as this one right away, without any further introduction: "{Kevtas/A kevta/The kevta/This kevta} wear(s) scarves." This was followed by the target image, which in turn was followed by the question used in the previous *kevtas* studies: "Which one of these is also a kevta?"

4.2.3.3 Procedure

The procedure was the same as in Section 4.2.2.

4.2.3.4 Analysis and Coding

The data of 12 participants were excluded because they either did not finish the task, self-reported not to be native speakers of English in the socio-linguistic questionnaire at the end of the study, had very fast or very slow response times. The remaining data of 398 participants were analysed and coded as in Experiment 8 (world knowledge).[5]

A 4 (wording: BP vs IS vs DS vs 'this', between participants) × 2 (connection type: PC vs SC, within participants) × 2 (block order: PC-first vs SC-first, between participants) mixed ANOVA was run to analyse whether the differences between the scores for generic subject types for each of the connection types were significant, and whether there were any significant differences based on which connection type was presented first. This analysis indicated a main effect of connection type ($F(1,396) = 16.238$, $p < .0001$), a main effect of wording ($F(3,396) = 107.831$, $p < .0001$), and a marginally significant main effect of block order ($F(1,396) =$

[5]Note that the data show the same pattern if the two participants who self-reported not to be native speakers of English, having grown up in an English-speaking environment until at least the age of 13 or whose parents did not both speak English to them at home. All the main effects and significant interactions are the same.

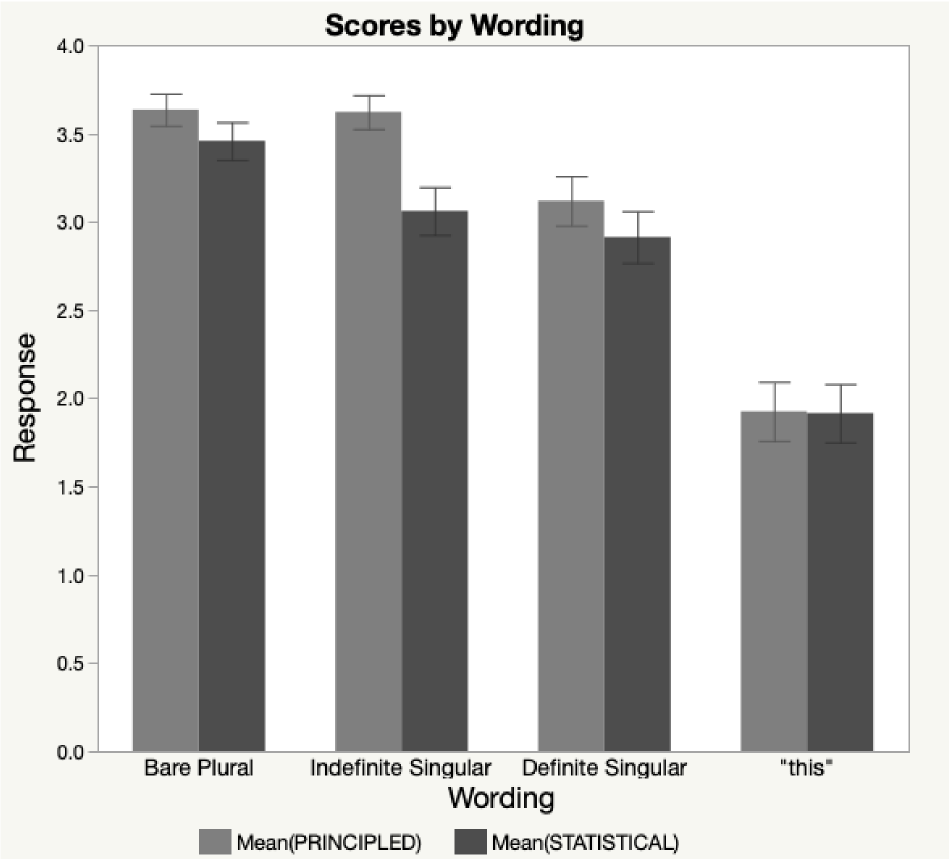


Figure 4.6: Averages of choosing property over shape matches for principled and statistical connections by wording. Error bars = ± 1 SE from the mean

10.336, p = .053). This analysis also indicated a marginally significant interaction of wording × connection type (F(1,396) = 2.433, p = .065), a significant interaction of connection type × block order (F(1,396) = 2.74, p = .033), a significant interaction of wording × block order (F(3,396) = 2.271, p = .08), and a significant interaction of wording × connection type × block order (F(3,396) = 2.738, p = .043). These results are visualised in Figure 4.6.

Planned comparisons for responses to trials with PC properties show that participants' responses showed differences between generic subjects. Participants were more likely to choose the *property* match if a PC property was trained with BP or IS subjects, compared to the DS (BP vs DS: p = .007; IS vs DS: p = .007). Paired comparisons for trials with SC properties show differences within generic subjects only for BP vs DS responses (p = .013), while BP and IS do not differ (p = .271) and neither do IS and DS responses (p = .177).

4.2.3.5 Results and Discussion

The results of Experiment 9 (world knowledge) (in particular the main effect of connection type, main effect of wording, and interaction of wording × connection type × block order) replicated findings in Hollander (2007); Hollander et al. (2009) and extended them to show that singular generic subjects are easily interpreted generically.

The repeatedly observed main effect of connection type could further indicate that principally connected properties are expected to be more generalisable than statistically connected ones.

Experiment 7 (world knowledge) and Experiment 8 (world knowledge) did not find the distinct asymmetries between IS vs BP and DS vs BP generic subject that were hypothesised in the literature review. This study showed a tendency towards the hypothesised interaction of wording and connection type, however only showed this at a significant level if block order is taken into account. One way to account for the lack of interaction could be attributed to the specific wording. The trials in Experiment 9 (world knowledge) did not have the initial prompt that Experiment 7 (world knowledge) and Experiment 8 (world knowledge) displayed. By removing the prompt, the BP that was the subject in this prompt was also removed. This is the crucial difference that set Experiment 9 (world knowledge) apart from the previous *kevtas* studies.

This is a subtle change and the results are closer to the hypotheses, the overall interpretation remains obscured as to whether they support the hypothesis of affecting distinct patterns based on morphosyntactically distinct subjects. One takeaway message from the first three *kevtas* studies is that subtle variation in the presence or absence of generic subjects, even if only presented in the initial prompts, can lead to distinct conceptualisations of novel kinds. This difference seems to particularly affect singular generic subjects, or those generic subjects that are less frequently used. Experiment 9 (world knowledge) demonstrated that the same stimuli and the same methodology can lead to different results, showing the hypothesised interaction for (between wording × connection type ×

block order) after not displaying the initial prompts which introduces the novel animal with a BP generic subject.

The next section of this chapter adapts these studies to make them more suitable for child participants. In order to translate the paradigm used for the adult studies in this section into a format in which young children can participate, all stimuli were recorded. These newly created audio files for each of the trials were added to the existing visual triads of images. The precise motivation and the ensuing studies are presented in the next section.

4.3 Audio-visual cues

To ensure a suitable adaption of the *kevtas* studies for child participants, audio-recordings were added to the existing, visual-only stimuli, used in Experiment 7 (world knowledge) to Experiment 9 (world knowledge), and tested in Experiment 10 (world knowledge) to Experiment 12 (world knowledge). While many studies in developmental psychology have experimenters read picture books to children, I decided to very carefully control the audio input that the child participants would receive. To keep the audio input consistent across participants, all sentences were recorded individually by the same native speaker, keeping in mind assumptions about the role of intonation, FOCUS, and pause in the interpretation of generic statements. I will illustrate the theoretical considerations of the stimuli construction here, particularly with regard to role of pause, lengthening, and pitch. I will present in detail the specific implementations for Experiment 10 (world knowledge), Experiment 11 (world knowledge), and Experiment 12 (world knowledge) in Section 4.4.1.2.

To recapitulate, the goal in the stimuli construction for all the *kevtas* studies, both on a verbal and now on an auditory level, is to guide participants towards a generic interpretation and away from an episodic one. One way of doing so it to add a pause after the subject DP, which focuses the noun in a way that enhances generic reference. Consider example (3-a), where the subject is kind-referring. Note how in this example, an intuitive way of pronouncing the sentence includes a lengthening of the subject and a slight pause before the predicate. Although semantically slightly odd, the minimal pair in (3-b) will be read without the same pause between 'dog' and 'is', and is likely to have emphasis on the demonstrative 'this'.

(3) a. The dog is a mammal.
    b. This dog is a mammal.

These observations are easily accessible via prosodic intuitions and have been discussed in similar and unrelated contexts. I will focus on intonation as one linguistic indicator of expressing FOCUS via prosody. While the stimuli in the experiments in this thesis all have preverbal subjects, there is similar research on the role of FOCUS intonation in post-verbal positions. Most notably in the context of interaction of intonation and information structure, Longobardi (2000, p. 694) argues that "in postverbal position the generic reading, whenever possible, always implies a detectable intonational break between the predicate and the postverbal subject, and a distinctly either 'given' (topic) or 'new' (focused) reading of the former. This is never required with the existential interpretation, which is perfectly compatible with a flat intonation and an internally unarticulated informational content of the sentence". Focusing further on prosodic considerations, phrase-final prosodic lengthening in English is manifested as a combination of extending the final syllable of the initial phrase, followed by a pause before the next word. In addition to that, phrase-final lengthening in generic subjects is accompanied by FOCUS pitch on the generic noun. Acknowledging this difference in constituent order, I considered the relevance and applicability of these observations for the current the stimuli design. Testing these minimal pairs, in terms of FOCUS prosody and pausing, on native speakers, it became apparent that this small (yet detectable) break is equally present for preverbal generic subjects.

One final piece of support comes from Kratzer & Selkirk (2007) who argue that a double pitch accent like "The **ele**phant lives in **A**frica" (where bold indicates pitch-accented syllables) gives two distinct pieces of novel information. Their line of thought is based on the idea that there are two spell-out phases, where 'the elephant' is the sentence's TOPIC and the predicate 'lives in Africa' also provides new information. With the arguments presented so far, this would be one further argument for phrase-final lengthening after the novel information/TOPIC provided by 'the elephant'. On the other hand, this then means that if something is de-accented, it is assumed that this is given information.

Therefore, all generic subjects were recorded with an emphasis on the subject noun itself and with a noticeable break before the predicate. Additionally, the noun was manually spliced into all three generic conditions such that the emphasis and pitch itself did not vary across generic conditions.

A related discussion concerns the phonological variant of the indefinite determiner 'a' that was used for the IS condition. English has (at least) two distinct variants for the indefinite determiner: the reduced or weak version that is prevalent in casual speech, pronounced \ə\, and the strong version, prominent in careful speech, pronounced as the diphthong \eɪ \. While they are often used interchangeably in everyday speech, based on a wide range of variables, only one version, the schwa, was used for the IS generic sentences. This decision was again made in order to keep the phonology as consistent as possible. Therefore, any differences between the IS generic items would not be due to the pronunciation of the indefinite determiner. This was crucial as it is one of the manipulated variables in this study.

Finally, there were considerations about the pronunciation of the baseline condition. The demonstrative determiner 'this' in the subject DP needs to both differ from the generic conditions and at the same time still sound natural. Therefore, the main intonation and focus on the subject DP fall on 'this' itself (illustrated in (4), emphasised in bold-face).

(4) **This** dog is a mammal.

While inserting FOCUS intonation on the noun and a pause before the predicate would make the sentence sound more similar to the ones in the generic conditions, it would also makes it sound unnatural. Moreover, it facilitates a subkind interpretation which is not intended here (compare the analogous example in (5) that easily allows for a subkind interpretation).

(5) **This** whale (namely the blue whale) weighs 176 tonnes.

Another option for the intonation of the baseline condition would be to simply adopt flat prosody. It was discarded for its various shortcomings: Not only does it sound less natural than adding a FOCUS intonation to some part of the sentence, it is also not engaging enough for young participants. Ultimately, having a baseline condition that is independently less likely to be preferred by the participants, as opposed to any of the more animated generic conditions, does not provide a true baseline for the critical conditions.

Based on these examples and the search for the most natural intonation, the best option for the emphasis is for it to be placed on 'this' itself.[6] Considering that wording is not a within-participants variable, it should not confuse participants but rather come across as a representative of natural speech.

4.4 Developmental interpretations of singular generic subjects

The focus on singular generic subjects is in itself a novel feature in match-to-sample paradigms. It was included to investigate the acquisition and use of generic statements in conceptual knowledge. Most match-to-sample studies have focused heavily on BP subjects. In light of the inclusion of child participants, a brief recap of the main assumptions regarding the development of singular generic subjects is warranted.

The main reason to collect developmental data for this research lies in its ability to provide us with a better insight as to the default of generic statements. Studies such as Leslie & Gelman (2012) have experimentally tested the developmental

[6]The information about the animal is, of course, also new, by virtue of it being a novel animal. The main accent could have also fallen on the noun for the novel animal itself. Compare "**This** bactra has three legs." with "This **bactra** has three legs." Both seem acceptable, but the first condition seems to convey the idea of a non-generic control condition better than the second one.

trajectory of generic sentences compared to quantified statements, albeit only of BP generics. While the recall rates for memory tasks for adults generally slightly favour generic statements over quantified ones, children falsely recall quantified statements as generic ones to a much higher degree (this is discussed in more detail in Sutherland et al., 2015 and Gelman et al., 2016). These studies look at what might be a 'default' form and are argued to provide insight into the conceptual system. The main conclusions that can be drawn from these investigations is that we as humans are biased towards generalising from one instance to a whole kind, and that this generalisation is significantly facilitated by generic language. They argue that even when we are presented with novel information in a majority-quantified form, we tend to store it as generic.

Cimpian et al. (2011) have similarly looked at developmental interpretation of IS generics. They highlight that English IS subjects are "typically used for individual referents that are either *nonspecific* or specific but *not presupposed* in the context of the conversation" (Cimpian et al., 2011, p. 1574) and that this interpretation emerges in children by the age of 3. They argue that the additional interpretation from nonspecific referents to *any*, i.e. to a generic referent, is extended by children further including TENSE and ASPECT marking into their computation. Their understanding thus develops to a more fine-grained representation of specific as opposed to generic reference by the age of 5 (and already starting at the age of 3). This could be interpreted as an argument in favour of *Generics-as-Default*, as in Cimpian et al. (2011) who interpret their findings on theoretical claims in Gelman & Raman (2003, p. 225). However, one might also argue that just because children develop interpretations from nonspecific to generic referents, this is insufficient evidence that generic interpretations are the cognitive default and that specific interpretations are more challenging. For instance, the supposed earlier emergence of generics could be due to the type of generic that is being investigated, as well as the type of specific construction against which it is contrasted. The empirical evidence in this thesis so far has not been able to reliably demonstrate that the fine-grained differences in the morphosyntax of generic subjects lead to behavioural differences. Yet, a conflation of the more fine-grained differences in these categories still glosses over some of the distinct semantic differences of the contexts in which they may appear. Bearing these opposing views in mind, I will return to an analysis of the control condition, with a specific referent introduced by the demonstrative determiner ‘this’, in the discussion of the studies at the end of this chapter.

Rhodes et al. (2012) are a notable exception to the heavy focus on BP generic subjects. They have also extended this paradigm by investigating social essentialist beliefs with both BP and IS generics, tested on both adult and child participants. They argue that the plurality of the generic subject is not the origin of essentialism, but that it is instead a property of generic subjects in general. This is a promising start for a more thorough analysis of the generic subjects available in English. Yet, there are three distinct differences in the methodology that ultimately warrant the design proposed in this section.

First, the inclusion of an IS subject in these studies was not sufficient. In their Experiment 2b, children were assigned one of three conditions (BP, IS, ‘this’) and their data were compared with those of adults, tested in one of two conditions (IS, ‘this’, Experiment 2a). They found no differences between the two generic conditions for the children. Note, however, that the probability of essentialist responses for both the BP and IS generics in this study were lower ($\tilde{3}3\%$) than those in the previous study (Experiment 1a and 1b for adults and children, respectively; where the child participants’ probability of essentialist responses was at $\tilde{4}5\%$ for BP generics). This cannot be due to a misinterpretation of the IS generic as episodic, though, as wording was manipulated as a between-participants variable. Most importantly, the difference between the specific control condition and the generic condition(s) was significant in both studies. Given what we know about the differential semantics, this paradigm, as well as the type of stimuli, might not be the right task to tease apart these differences. Moreover, while the inclusion of a singular generic subject is a step towards an investigation of the full paradigm, the semantically distinct kind-referring DS subject is still lacking. This is especially important when considering that the BP is distributed more widely than the IS and DS. A lack of differences between the BP and the IS in these studies is not insightful in terms of where the differences come from.

Second, their methodology relies to a large extent on explicit memory recall. It is questionable, however, whether the form that is explicitly recalled in memory tasks can be equated with underlying conceptualisation of a kind's properties (for further discussion, see especially Lazaridou-Chatzigoga et al., 2015, but also Leslie & Gelman, 2012; Sutherland et al., 2015). In fact, Chapter 3 has argued that the overt form is not necessary but that expectations with regard to a kind's make-up are raised by subtle cues which are potentially not readily available to explicit memory (and this will be further seen in the role of participant comments in the upcoming studies as well as the study in Chapter 5). Further, the heavy reliance on both English stimuli and BP generics means that most studies are run with generic subjects that are morphologically bare and have fewer overt pieces than their quantified (or specific) control counterparts. To address such a bias in favour of English, Gelman et al. (2016) have compared these recall rates found in English with Spanish stimuli. Spanish uses definite plural subjects (similar to French in Section 3.6) and therefore looks more complex than a BP. Additionally, it also bears more resemblance to the composition of the control conditions in these studies, be they specific or quantified subjects. In reference to the French pseudoword studies, recall that the results for the expectations of the category-property links between English and French were distinct and that they can therefore not be treated as direct equivalents in terms of their underlying representations.

Third, the research question of this thesis is not specifically targeted at the notion of essentialism. Rather, I investigate the semantic properties of generic subjects, regardless of their category membership, and focus more on the expectations of category-property links based on the morphosyntax. While we admittedly rarely talk about the essentialism of artefacts, essentialism is a much more meaningful metric to assess the representations of novel animal kinds. The experiments below still use animal kinds, but the hypotheses for the distinct linguistic conditions have been formed based on the semantics of the morphosyntactic pieces, regardless of category membership.

The age-groups for the child participants were set up in a way that allows us to look at the effect of entering formal education. Recall that an analysis of the CHILDES corpus data in Chapter 2 showed that children are virtually unexposed to IS and DS generic subjects, while BP subjects are ubiquitous. If children are able to interpret singular subjects, definite and indefinite, early on in their lives, this will provide more evidence for the hypothesis that generic subjects, regardless of their specific form, are a default way of conceptualising world knowledge. If, on the other hand, there is a developmental trajectory with clearly distinct patterns for either each of the generic subjects, or, more generally, for plural vs singular subject forms, this provides further evidence for the hypothesis that they are learned, and learned in a formal educational setting.

To ensure that the audio stimuli provide data similar to the previous studies, Section 4.4.1 starts with a replication of those studies. Section 4.4.2 presents a version of this study run in a lab setting before Section 4.4.3 presents the data collected from child participants.

4.4.1 Experiment 10 (world knowledge): Adult participants

To ensure that this version of the study, with the new audio stimuli, is comparable to Experiment 9 (world knowledge), adult speakers were tested first.

Just as for the visual-only stimuli, a baseline test was run with the audio stimuli (in-person in the Conceptual Development Lab). Due to the number of previous studies with the same visual stimuli, this was a small study. Adult native speakers of English (n=27) were recruited to respond to audio-recorded sentence of the form *This is a fep.* annd received course credit for their participation. They then chose one of the two options, the *shape* or the *property* one. Results were in line with the previous baseline tests and are presented in Table 4.1.[7]

[7]Note that the *shape* scores for the vorzyd are a lot higher than the other ones. Given the other baseline test that was run, this does not give rise to immediate concern. In fact, if anything, participants that choose the property option overcome a bigger *shape* bias than for the other critters.

Table 4.1: Baseline judgements for adult audio stimuli

	% shape	% property
Bactra	0.85	0.15
Scred	0.7	0.3
Fep	0.78	0.22
Febbit	0.63	0.37
Vorzyd	0.96	0.04
Plog	0.74	0.26
Bant	0.81	0.19
Kevta	0.67	0.33

4.4.1.1 Participants

434 adult native speakers of American English were recruited via Amazon's MTurk and redirected to the survey, which was hosted on Qualtrics. None of them had previously participated in any of the earlier versions of the *kevtas* studies.

4.4.1.2 Stimuli

The stimuli were the same as in Experiment 7 (world knowledge) through Experiment 9 (world knowledge) in Section 4.2. There was a minor change to one of the previously used properties, such that *goes to the amusement park* was changed to *goes to amusement parks*. This was judged to sound more generic and not as if the animal in question goes to one particular amusement park. While this is noticeable, but not as prevalent, for the BP condition (e.g. the analogous *Children go to {the amusement park/amusement parks}.*), it is more noticeable for the singular conditions (e.g. *{The/A} child goes to {the amusement park/amusement parks}.*). Note that the statistical property, i.e. *going to the amusement park*, is not supposed to give a generic interpretation when combined with a definite singular.[8] This is analogous to *The car has a radio*, which is interpreted as episodic and not generic,

[8] The property of going to the *amusement park* was chosen specifically because it has a stronger artefact connotation than just *the park*. Thinking about some of the novel animal kinds, e.g. *vorzyds*, which look like butterflies, it holds that they could go to the park because of the flowers there, for example. This should not be the first association for an amusement park, however, where the main attractions are rides. Interestingly, the different effects for the definite and indefinite predicate do not seem to be as clear for the property of going *to the park*, as in *Children go to the park* and *Children go to parks*.

as opposed to the BP, *Cars have radios*, which allows for generic interpretations. The underlying justification to make the stimuli as generic as possible this is that in order to find variation within the generic interpretations, all three generic phrases need to receive a generic interpretation. This effort has been made in previous studies and was considered in the design of the audio files. As such, it is not a departure from the design but rather an effort to be more consistent across the various conditions. No other changes were made and the study remained as a 2 (connection type) × 4 (wording) design.

Since these are the first stimuli of this type of experiment that are audio-recorded, various decisions were made about the phonological properties of the stimuli, as explained in Section 4.3. The FOCUS intonation for the generic subjects was manifested by a change in pitch for the novel critter, a lengthening of the critter as it coincided with the being the phrase-final constituent, and a pause between the critter and the predicate. Having included a FOCUS intonation that indicates that the subject is new, and not given, is an extension of Longobardi's observations of generic interpretations. The pause was matched between the IS and DS, by using the same recording for both conditions, so that as many variables as possible are kept constant between the singular generic conditions. This was extended to the rest of the stimuli whenever possible. To elaborate on the examples in (3) above (repeated here as (6)), this means that the pronunciation is explicitly matching the following pattern, applied to the stimuli for this study in (7):

(6) The **dog** $_{\text{pause}}$ is a mammal.

(7) The **kevta** $_{\text{pause}}$ has curly fur.

An illustration of how the overlap was maximised between conditions is shown in the colour-coded graphic in Figure 4.7, where the same colour indicates the same recording being used across multiple conditions (and which additionally includes numbers in parentheses indicating the same recordings). This allows an investigation of effects between conditions that is much more likely due to morphosyntactic rather than phonological differences.

A	(1)	Bactra	(1)	Collects	(2)	Plastic rings	(2)
The	(7)	Bactra	(1)	Collects	(2)	Plastic rings	(2)
		Bactras	(4)	Collect	(4)	Plastic rings	(2)
This	(5)	Bactra	(5)	Collects	(2)	Plastic rings	(2)
This	(6)	Bactra	(6)	Collects	(2)	Plastic rings	(2)
A	(1)	Bactra	(1)	Has	(3)	Three legs	(3)
The	(8)	Bactra	(1)	Has	(3)	Three legs	(3)
		Bactras	(4)	Have	(8)	Three legs	(3)
This	(5)	Bactra	(5)	Has	(3)	Three legs	(3)
This	(6)	Bactra	(6)	Has	(3)	Three legs	(3)

Figure 4.7: Experiment 10 (world knowledge): Table illustrating overlap of audio stimuli for *kevtas* studies, where the same colour and number indicate the same source audio-file.

Recall from the discussion in Section 4.3 that the allophone chosen for the indefinite determiner was the weak version, expressed by a schwa (\ə\) as opposed to the strong allophone found e.g. in FOCUS position and/or to emphasise contrast (\eɪ\). This decision was informed by native speaker intuitions (for both British and American English speakers), declaring that the emphasised version of the indefinite determiner, \eɪ\, sounds more episodic and less generic.[9] Each recorded sentence was inserted into the study together with the visual cues, i.e. a *target* image, a *shape* match, and a *property* match.

The final difference that was implemented was a resizing of the images. While in the previous versions of the *kevtas* studies, participants saw a large version of the target image and then scrolled down to see the two options side-by-side, the images were scaled down such that all three pictures could be seen simultaneously. This should make it easier, particularly for child participants in Experiment 12 (world knowledge), to compare the target image to both answer options at the same time.

4.4.1.3 Procedure

The procedure was similar to the one in the previous *kevtas* studies. In this version, participants were able to replay the stimulus if they clicked the button for the audio

[9] Why exactly this phonologically induced difference in interpretation exists in this scenario is not clear to me at the moment.

file more than once. However, they were not explicitly instructed or encouraged to do so by the experimenter or in the instructions.

All stimuli were presented completely randomised and not in blocked order. This was to test whether the two blocks of four animals had an effect on the interpretation.

After answering the eight trials, they were presented with an essentialism questionnaire and a sociolinguistic background questionnaire (both of which were optional for participants), the same ones as in the previous experiments.[10]

4.4.1.4 Analysis and coding

Participants who self-reported that English was not their native language in the optional sociolinguistic questionnaire were excluded from the analysis (N = 5). The analysis of the data is based on the remaining 429 participants.

Results were coded in line with previous experiments, i.e. answers for the *shape* option were coded as “0”, *property* answers were coded as “1”, indicating that participants generalised the predicated property and that their generic interpretation over-wrote the assumed *shape*-bias. Every participant saw eight trials, four of which expressed principally connected properties and four expressed statistically connected properties. As in Experiment 8 (world knowledge) and Experiment 9 (world knowledge), participants could therefore score a maximum of score of 4 for PCs and 4 for SCs.

A 4 (wording: BP vs IS vs DS vs ‘this’, between participants) × 2 (connection type: PC vs SC, between participants) ANOVA was conducted over scores of *property* > *shape* response options to establish whether the differences between the scores for generic subject types for each of the connection types were significant. This analysis indicated a main effect of connection type ($F(1,425) = 15.399$ $p < .0001$), and a main effect of wording ($F(3,425) = 15.767$, $p < .0001$). There was a marginal interaction of wording and connection type ($F(3,425) = 2.502$, $p = .059$). These results are visualised in Figure 4.8.

[10] Refer to Appendix C for the questionnaires.

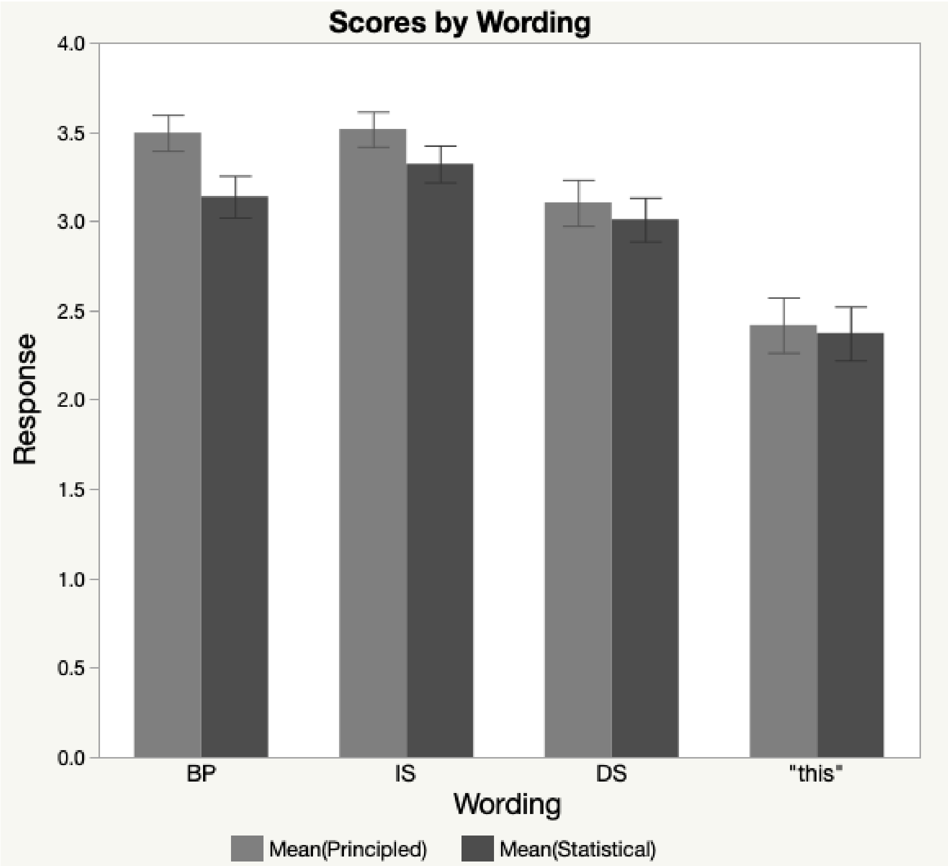


Figure 4.8: Averages of choosing property over shape matches of adult listeners by wording and connection type. Error bars = ± 1 SE from the mean

Planned comparisons show that within PCs, all generic subjects differ from the control condition (BP and IS vs 'this': $p < .0001$; DS vs 'this': $p = .0005$). The data for pairwise comparisons of SCs also show that all generic subjects are different from the control condition (BP and IS vs 'this': $p < .0001$; DS vs 'this': $p = .0019$), but no differences within generic subjects (all p values between .31 - .89).

4.4.1.5 Results and discussion

The same general trends were found in these experiments (a main effect of wording and of connection type, a marginally significant interaction of wording × connection type), albeit at a slightly weaker level. This could be due to the randomised presentation of the properties, rather than presenting them in blocked order as before.

This study has shown a marginally significant interaction between the wording of the subject and connection type. This could indicate supporting evidence suggests that the IS enhances expectations for PCs over SCs. However, the strongest pattern was found for the BP, which was not hypothesised to show differences between PCs and SCs. Further, the DS demonstrates similar patterns, but at a lower magnitude. Even so, the DS data look sufficiently different from the control condition. The novel animals introduced with ‘this’ do not differentiate between the two connection types, such that this main effect is driven by the generic subjects. The differences for PCs and SCs, compared between the BP, IS, and DS, look mostly similar, unlike the hypothesised differences. The biggest contrast between this hypothesis and the data is the high acceptance of SCs for DS subjects.

The purpose of this iteration of the *kevtas* studies was to test whether the audio-recorded stimuli would yield similar results to the visual-only stimuli. Since the results were comparable to the previous studies with adult participants, they seem appropriate for the child studies. The randomisation of the critters was added to test the role of blocked order for connection type presentation. The effects were replicated, but overall they were stronger in the blocked design. Before moving on to the study with child participants (Experiment 12 (world knowledge)), I will briefly present Experiment 11 (world knowledge) in the next section. This version of the study recruited adult participants, but these were tested in a laboratory setting and with a blocked design. This version provides the closest match with the conditions under which the child participants were tested.

4.4.2 Experiment 11 (world knowledge): Adult participants (in lab)

This experiment tested the audio-visual stimuli presented in Experiment 10 (world knowledge). This version was run so as to test the study in the same conditions in which children will be tested, i.e. with a blocked design and in a lab.

4.4.2.1 Participants

96 adult native speakers were recruited either via the subject pool of a large university or when coming to the lab as caregivers to child participants. None of them reported that they had participated in an online version of this study

before. They were compensated either in course credit (if student participants) or received \$10 and free parking for participation.

4.4.2.2 Stimuli

The stimuli were the same as in Experiment 10 (world knowledge).

4.4.2.3 Procedure

Participants were asked to read the Participant Information Sheet and sign the consent form. The procedure and components of the experiment itself were the same as in Experiment 10 (world knowledge), with the only difference being that they used the blocked design from Experiment 8 (world knowledge) and Experiment 9 (world knowledge).

4.4.2.4 Analysis and coding

The data of all 96 participants were analysed. The data were coded as they were in Experiment 9 (world knowledge), with the additional variable of order, resulting in a 2 (connection type: PC vs SC) × 4 (wording: BP vs IS vs DS vs 'this') × 2 (block order: PC-first vs SC-first) mixed design.

A 4 (wording: BP vs IS vs DS vs 'this', between participants) × 2 (connection type: PC vs SC, within participants) × 2 (block order: PC-first vs SC-first, between participants) mixed ANOVA was run to analyse whether the differences between the scores for generic subject types for each of the connection types were significant, and whether there were any significant differences based on which connection type was presented first. This analysis indicated a main effect of connection type ($F(1,88) = 8.253$, $p = .005$), a main effect of wording ($F(3,88) = 26.755$, $p < .0001$), no main effect of block order ($F(1,88) = .413$, $p = .522$). This analysis indicated that there was no interaction of wording × connection type ($F(3,88) = .753$, $p = .524$), but that there was a significant interaction of connection type × block order ($F(1,88) = 14.381$, $p < .0001$), a significant interaction of wording × block order ($F(3,88) = 3.282$, $p = .025$), and a significant interaction of wording × connection type × block order ($F(3,88) = 20.261$, $p < .0001$). These results are visualised in Figure 4.9.

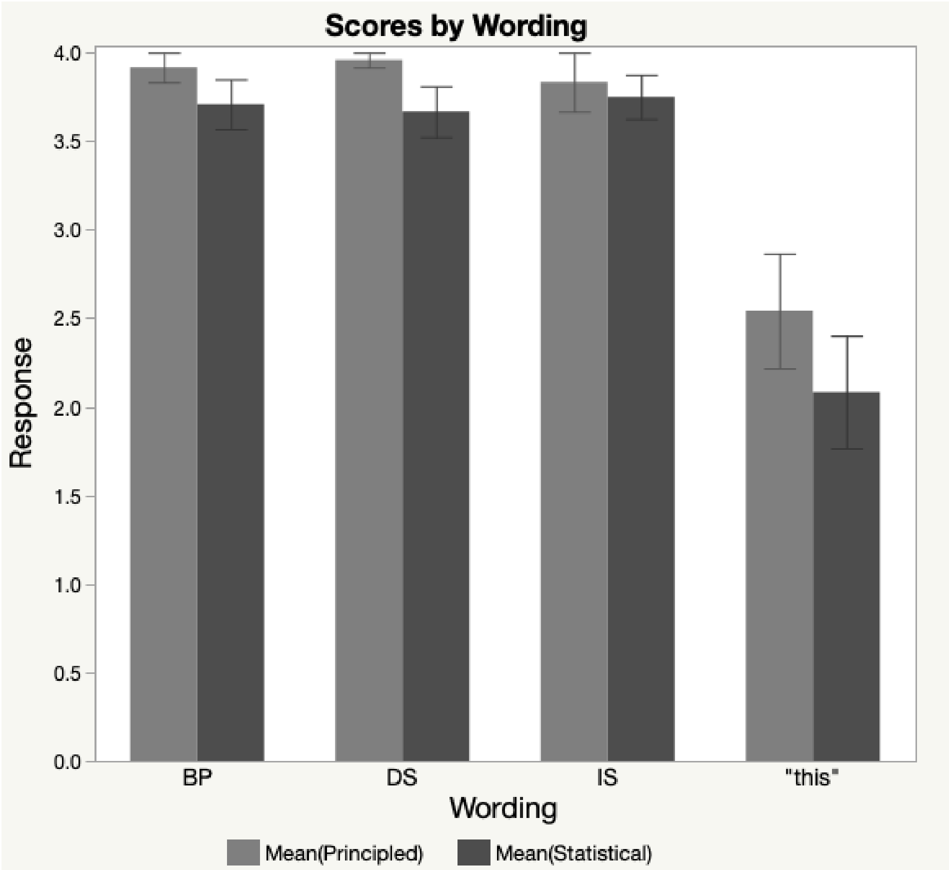


Figure 4.9: Averages of choosing property over shape matches for principled and statistical connections by wording. Error bars = ± 1 SE from the mean

Planned comparisons show that within PCs, all generic subjects differ from the control condition (all p values $< .0001$), as was the case for Experiment 10 (world knowledge). The data for pairwise comparisons of SCs show the same pattern, where all generic subjects are different from the control condition (all p values $< .0001$), but no differences within generic subjects.

4.4.2.5 Results and discussion

The results (a main effect of wording, connection type, and an interaction of wording $\times$ connection type $\times$ block order, but not of wording $\times$ connection type by itself) replicate the findings of the previous adult studies.

The analysis of order also showed that, in BP and DS wording conditions, participants are more likely to choose the *property* option, indicating a generalisation of the property as opposed to choosing the visually more similar *shape* option, if they were first shown statistically connected properties. The opposite trend was observed for the baseline condition.

Overall, it seems useful, and justifiable, to keep the blocked order for the child studies. It seems advantageous to employ an experimental design that renders stronger effects, seeing that child data can be noisier than adult data. This is especially important since our primary interest for the child studies lies in finding an account for a developmental trajectory.

The next section will present the experiments and results for the children participating in this study.

4.4.3 Experiment 12 (world knowledge): Child participants

With regard to the acquisition of singular generic subjects in Section 4.4 above, it is important to bear in mind that the interpretation of kind-referring DS subjects has not yet been tested with child participants. Therefore, little is known about whether children default to a generic interpretation in these match-to-sample tasks, or whether they will mostly interpret them episodically, similar to the control condition. Based on Rhodes et al. (2012), children could conceivably interpret DS subjects like they do IS subjects, and generalise to the same extent as they do with BP subjects. However, there are two key differences between these experiments: First, the task type itself is different as Rhodes et al. (2012) used a picture book with a narrative based around one single animal. Second, they measured essentialist responses based on a battery of different questions. Instead, this study aims to provide a more direct measure of whether properties are generalised, based on the competition between *shape* and *property* matches for the target animal.[11]

4.4.3.1 Participants

297 children between the ages of 4 and 10 years participated in the study in the Conceptual Development Lab at the University of Michigan or at the Ann Arbor Hands-On Museum. They were split into three age groups: 4-5 years, 6-7 years,

[11] Note that another baseline test was run (n=72) in either the Conceptual Development Lab or the Living Labs. Results were in line with previous baseline tests.

and 8-10 year, with n ≈ 100 per age group, and n ≈ 75 per wording condition. This large number of participants is due to the 2×4×3 design (detailed in Section 4.4.3.2 below), which provides approximately 12 participants per condition. Families were contacted through the lab's database of families that have previously expressed interest in participating in on-campus studies or were invited to participate when visiting the museum.

4.4.3.2 Stimuli

The stimuli are the same as those in the experiment with adult participants above (Section 4.4.1). Participants were matched to one of three age groups. This resulted in a 2 (connection type: PC, SC) × 4 (wording: BP, IS, DS, 'this') × 3 design (age group: 4-5, 6-7, 8-10 years).

To make sure that the conditions were equally split between the three age groups of the participants, the stimuli were separated into sets that were chosen according to the age of the participant. Half of the eight stimuli were presented with their principally connected properties, and half with their statistically connected properties, as illustrated in Table 4.2.

Table 4.2: Conditions and grouping of stimuli for child participants

SET #	SUBJECT FORM	PC	SC
1	BP	Group A	Group B
2	BP	Group B	Group A
3	DS	Group A	Group B
4	DS	Group B	Group A
5	IS	Group A	Group B
6	IS	Group B	Group A
7	Control	Group A	Group B
8	Control	Group B	Group A

The stimuli in Group A consist of bactras, screds, feps, and febbits. Those in Group B are vorzyds, plogs, bants, and kevtas. This ensures, in line with the previous experiments, a balance of aquatic, air-, and land-dwelling animals as well as an even spread of mono- and bisyllabic names across the groups.

The order for each group is randomised in every trial. Further, the order of items is random, such that all eight items are randomly intermixed.

4.4.3.3 Procedure

The procedure is similar to the adult version above. Any differences are due to the fact that the child participants were guided through the experiment by an experimenter, as opposed to adults who participated on their own.

Upon arrival in the lab, parents provide consent for child participation and children provide oral assent to participate. When children participated in the museum, they found the experimenters in an area designated to the Living Lab, but the overall procedure was the same.

At the beginning of the study, a pre-recorded statement ("Let's get started.") was played before any of the testing stimuli to orient the child to the tablet and ensure the equipment worked well and the audio was at an appropriate level. This was recorded by the same person as the stimuli, so that the child could get habituated to the voice as well.

During the testing phase, the child was presented with 8 trials, consisting of the stimuli triples identical to those used in Experiment 4.4.1. As with the adult study, each recorded sentence was presented together with the visual cues, i.e. a *target* image, a *shape* match, and a *property* match. The audio stimulus was played to the child twice, unless the child started to carry out the task after hearing the sentence the first time.[12]

They were then asked a question of the form "Which one of these is also a *plog*?" (where *plog* stands in for any of the novel animals) and the experimenter asked "This one? Or this one?". While doing so, they looked and pointed at each of the two options, going from left to right, and without lingering over the second picture. The presentation of the *shape* and *property* match was random, so that the experimenter would not always point at one or the other answer choice first.

[12] The stimuli were not automatically played twice. The reasoning for this is that after a couple of trials, the child is quite likely to understand the task and might not need to hear the statement a second time. In these cases, it might be perceived as pragmatically odd to repeat the phrase while the child is already answering the question.

If the child answered right away, the experimenter would not continue pointing and asking so as to have the setting matching conversational turn-taking as pragmatically natural as possible. This is in line with not playing the audio-recording twice if the child starts to answer the question after hearing the question the first time.

4.4.3.4 Analysis and coding

The data of 297 child participants were coded as in the previous experiments. They were analysed according to the variables *connection type*, *wording*, and *age group*.[13]

A 4 (wording: BP vs IS vs DS vs 'this', between participants) × 2 (connection type: PC vs SC, within participants) × 3 (age group: 4-5 years vs 6-7 years vs 8-10 years, between participants) mixed ANOVA was run to analyse whether the differences between the scores for generic subject types for each of the connection types were significant, and whether there were any significant differences based on the age group of the participants. This analysis indicated a main effect of connection type ($F(1,285) = 23.259$, $p < .0001$), no main effect of wording ($F(3,285) = .581$, $p = .628$), but a main effect of age group ($F(2,285) = 27.931$, $p < .0001$). This analysis indicated that there was no interaction of wording × connection type ($F(3,285) = 1.735$, $p = .160$), but that there was a significant interaction of wording × age group ($F(6,285) = 2.267$, $p = .037$). There was no significant interaction of connection type × age group ($F(2,285) = .398$, $p = .672$) and neither a significant interaction of wording × connection type × age group ($F(6,285) = .334$, $p = .919$). These results are visualised in Figure 4.10.

Pairwise comparisons show different developmental patterns based on the wording condition. Only the 8-10 year olds showed this interaction of wording by connection type at a significant level, and scores for IS and DS subjects were greater than for the control condition 'this'. The scores for BP subjects show a non-significant trend in the same direction. Note that in general, the proportion of *property* vs

[13] An initial analysis including the variable 'block order' did not show any order effects. Therefore, the following analysis is based on using only differentiating the data based on these three variable categories.

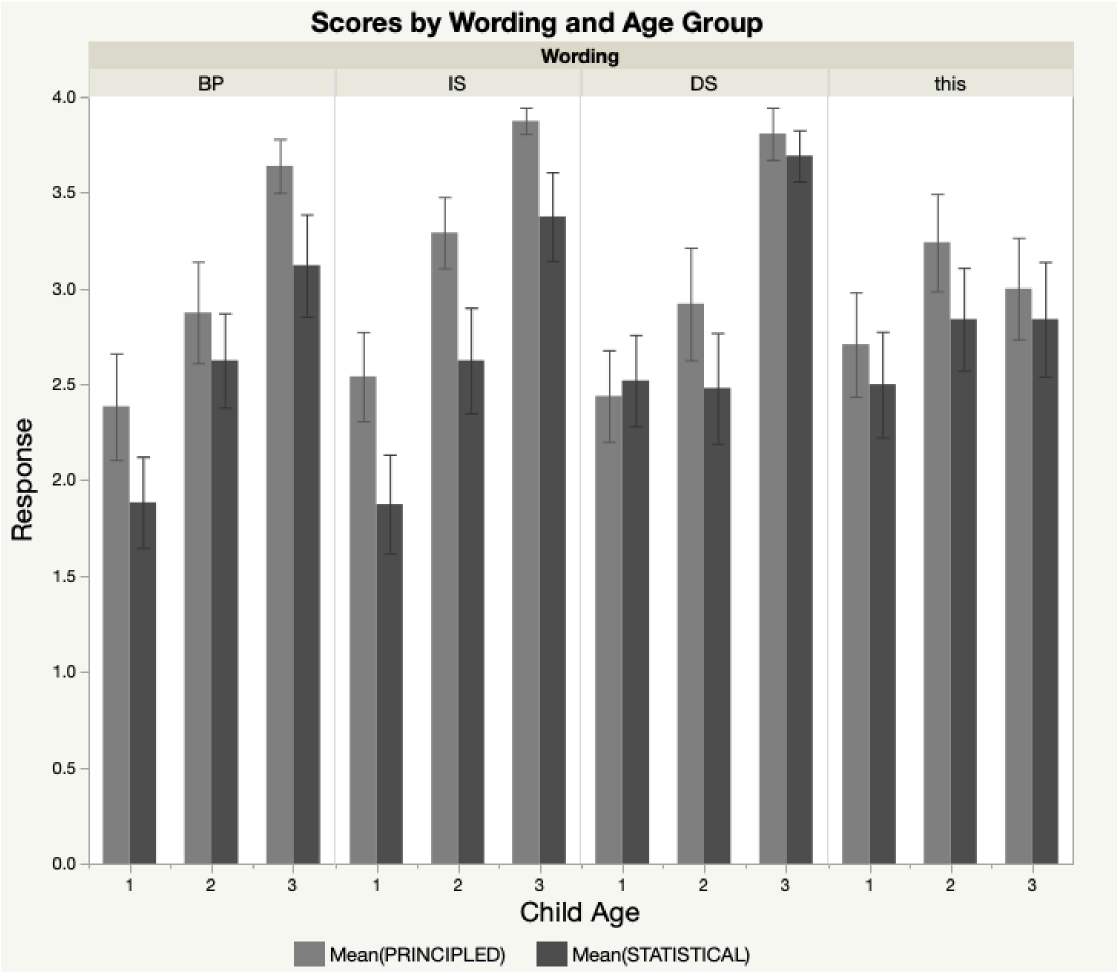


Figure 4.10: Averages of choosing property over shape matches of child participants by age group, where 1 = 4-5 year olds, 2 = 6-7 year olds, and 3 = 8-10 year olds. Error bars = ± 1 SE from the mean

shape choices for the control condition was higher than the one in Experiment 10 (world knowledge), with adult participants. This might explain the trending towards, but not at significance level difference between BP and 'this' subjects. Recall that Hollander et al. (2009) found an overall tendency for children to generalise, which could explain the higher scores for the control condition compared to those given by adult participants.[14] Note further that for the child participants, the scores for PCs in the control condition are almost on par with those for DS subjects.

Focusing on the developmental findings within this study, for both BP and IS sub-

[14]To recall a previous statement, I will look into potential task-dependent effects and methodological biases in Chapter 6 and provide an alternative explanation as to why the baseline scores are high, and why *shape* is potentially not the best baseline for *property* in this experimental set-up.

jects, the proportion of choices for *property* as opposed to *shape* responses increased with every age group. By contrast, the scores for DS subjects only increased between the 6-7 and 8-10 year-olds. For the baseline condition, there were no differences. The mean scores for *property* over *shape* scores are illustrated in Table 4.3.

Table 4.3: Developmental patterns for proportion of *property* over *shape* choices. Symbols indicate significant differences between age groups.

	4-5 years		6-7 years		8-10 years	grand mean
BP	2.135	$<$	2.813	$<$	3.38	2.776
IS	2.208	$<$	2.958	$<$	3.625	2.930
DS	2.48	$=$	2.7	$<$	3.75	2.977
'this'	2.604	$=$	3.04	$=$	2.92	2.855
grand mean	2.357		2.878		3.419	

Even though the data did not show an interaction of wording × connection type across the board, given that this was the main hypothesised interaction for this experiment, pairwise comparisons for this interaction were computed. These show the hypothesised interaction of wording × connection type, but only for BP ($p = .003$) and IS subjects ($p < .0001$), where the proportion of *property* over *shape* choices was significantly higher. The were no differences for DS subjects ($p = .3$) or 'this' ($p = .099$). These effects are illustrated in Figure 4.11.

4.4.3.5 Results and discussion

The developmental changes that are revealed by the data, in particular those between the different types of generic subjects summarised in Table 4.3, are, to an extent, consistent with the idea that DS generic subjects are acquired after BP and IS generic subjects. However, all results show that adult-like results are only given by the oldest age-group. In fact, the proportion of *shape* over *property* choices are highest for the control condition for both 4-5 and 6-7 year old participants. This pattern is only reversed in the 8-10 year olds. Therefore, these data do not provide enough information to generalise to overall developmental patterns. It could equally be argued that the DS is acquired at the same time as the other two generic subjects. This is particularly defensible seeing that the results for the DS of

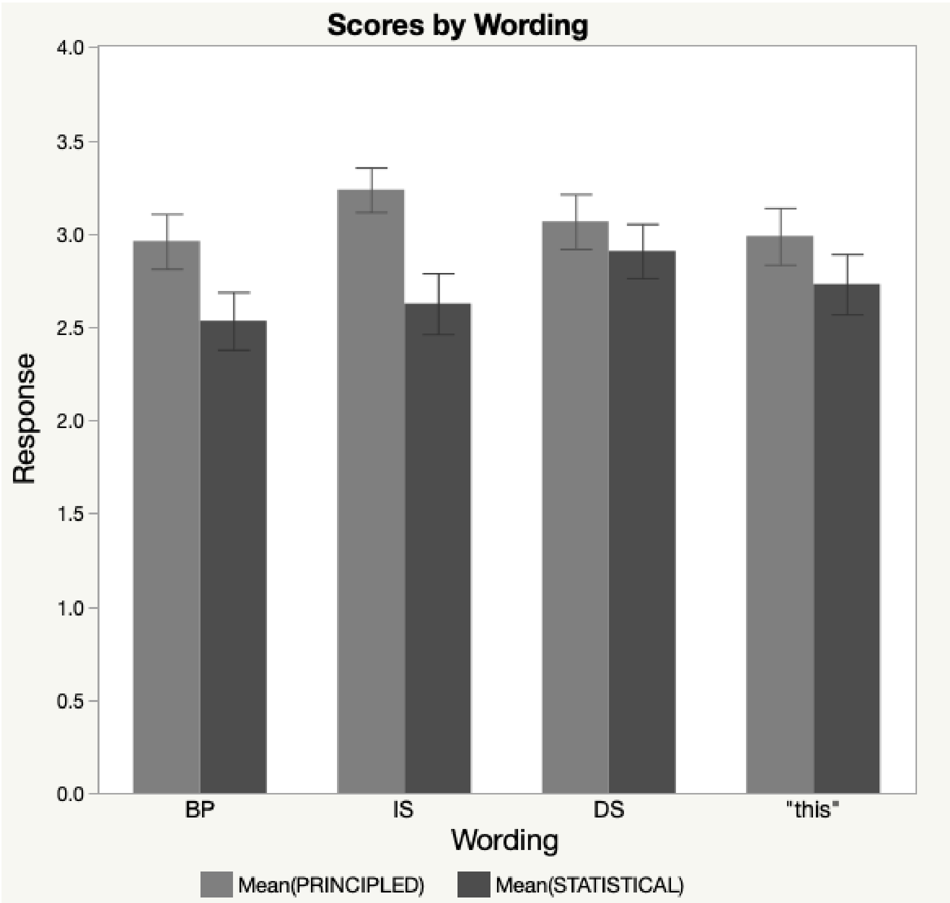


Figure 4.11: Averages of choosing property over shape matches of child participants by wording and connection type. Error bars = ± 1 SE from the mean

the youngest age group are comparable to the results of the BP and IS, and again for the oldest age group. The biggest deviation between the DS on the one hand, and the BP and IS on the other hand, is in the middle age group. Future studies should investigate developmental patterns in more detail to determine whether the DS indeed is acquired after the BP and IS.

Should these and future results indicate that the three generic subjects have distinct developmental patterns, with scores for DS subjects indicating a later acquisition age, then that interpretation also needs to account for the high scores for 'this'. One explanation for a hypothesises later acquisition of DS generic subject might lie in the role of formal education. These formal settings likely expose children to a higher frequency of these more restricted kind-referring subjects. While the acquisition of generic interpretations for DS subjects might be supported through exposure and frequency, the underlying reason for the acquisition could still be due to an innate capacity for generic DS interpretations, in a similar fashion to the acquisition of BP and IS generic interpretations.

Returning to the observed interaction of condition × wording for IS and BP subjects, this is in line with the prediction that any differences would be at a higher magnitude for IS compared to DS subjects, given their prevalence and linguistic restrictions. Further, the stronger interaction for IS compared to BP subjects is in line with the hypothesis that the semantics of IS subjects support PCs more than SCs, while BP subjects may support either type of generalisation.

Overall, based on these results alone, it is not possible to decide whether this means that generics are acquired later, and in this case after specific interpretations, or whether the demonstrative ‘this’ is, at a young age, also interpreted as making generic reference. However, the literature on the acquisition of generics, where only the BP and ‘this’ are compared, points strongly towards a default generic and a later acquisition of specific interpretation.

In general, the *kevtas* studies provide only limited evidence for the hypothesis that morphosyntax is indeed used to distinguish between different types of generalisations. Interestingly, children already interpret singular subjects generically in this task, and at the same time, they are better able to use morphosyntactic cues as they get older. Considering why the scores in favour of generic interpretations were high across the board, one possible explanation may lie in the theory that the denotation of nouns as names for kinds always evokes a kind, as soon as a label in the form of a noun is provided Mueller-Reichau (2013). Alternative support comes from *Generics-as-Default*, proposed by Leslie & Gelman (2012), arguing that in the absence of evidence to the contrary, children (and adults) have no reason to assume that they should not generalise based on the, possibly only very few, instances to which they have been exposed. Both theories have been discussed throughout and could explain the results found in the experiments of this thesis so far.

Linked to this possible explanation is the idea that the category type itself affects the perception of category-property links. Animal categories are known for their high essentialism and might therefore evoke a tendency to support both *property* and *shape* matches. Incidentally, this was part of the reason Hollander (2007) initially decided on animal kinds as opposed to artefacts. In addition to their high essentialism, or potentially precisely because of it, Hollander reports that generic language is reported to be more common when talking about animal kinds (Gelman et al., 1998; Gelman & Tardif, 1998; Goldin-Meadow et al., 2005). It might be easy for child participants to generalise to the whole kind from these images, potentially even disregarding the linguistic form to a larger extent than if no pictures had been included as visual aids. This could still explain the consistent main effect of connection type, since PCs are more easily generalised for animal kinds than SCs.

Finally, the relatively high proportions of *property* over *shape* matches for IS subjects that express SCs need to be addressed because normally, IS generics do not support SCs. One possible explanation might be that participants have provided a "charitable interpretation" of the stimuli and the task. The formulaic set-up of the stimuli and the repetition through the study make them sound very instructional. This hypothesis was not specifically tested but could be addressed by changing the task type. I will return to this question in Experiment 14 (production) and Experiment 15 (production) in Chapter 6.

The results need to be further evaluated in light of the higher scores for PCs even within BP subjects, given that they are hypothesised to support PCs and SCs equally. I will address the repeatedly found main effect of connection type, favouring PCs over SCs across all three generic subject types, in the next chapter when raising the issues of overhypotheses.

4.5 General discussion

This chapter presented the results of six studies. They were in novel in their systematic comparison of BP, IS, and DS subject forms. Previous versions of these match-to-sample tasks have investigated genericity by using only BP subjects. The studies in this chapter have shed light on the accessibility of generic interpretations beyond the commonly studied BP. It has further taken the insight from the corpus analysis in Section 2.8, which showed that linguistic input of singular generic subjects is virtually absent in parent-child interactions, and investigated whether this means that singular generics are therefore only acquired later, and possibly only in an instructional setting such as is provided via formal education.

Parallel to the experimental results in Chapter 3, these studies did not reliably show the hypothesised predictions. The main findings are as follows:

- All studies showed a main effect of connection type, consistently showing higher scores for PCs over SCs. This includes a preference of PCs over SCs for BP subjects, unlike the hypothesis in the literature predicting that BPs are equally suitable for PCs and SCs. I will address other possible reasons for this in the following chapters.

- A main effect of wording was prevalent across the studies, showing that generic subjects differ from the specific baseline 'this'.

- There were only a few differences within generic subjects, but overall, generic subjects behaved similarly across experiments.

- The hypothesised interaction between wording and connection type was only found in some of the studies. Crucially, after removing the initial prompt, *Do you know about kevtas?*, which includes a BP subject, the interactions were more reliably found. In general, however, IS subjects did not reliably show higher scores for PCs over SCs, compared to BP subjects.

The latter three studies have relied on audio-stimuli in an attempt to remove variation in the auditory make up of the stimuli between participants and experimenters. One shortcoming that became apparent only during data analysis and interpretation, with regard to the high scores across the board, is that the audio files might be perceived more like formal instructions, at least by the children. This could be remedied by including an additional, conversational-sounding sentence, such as "That's what I know about {kevtas/a kevta/the kevta/this kevta.}" In this case, the morphosyntax would again match the wording condition and the intonation would be based on the original considerations regarding FOCUS as well.[15]

The findings of these studies did not reliably show the hypothesised interaction between wording and property. Notably, the effect of removing the initial prompt showed more consistent effects of overall interactions between wording × connection type × block order, yet not reliably between the factors of wording and connection type alone. The next chapter will focus on a more natural learning setting, provided via narrative learning. Recall that some previous learning studies targeting the acquisition of genericity have used picture book studies. Most notably, Gelman et al. (2010) and Rhodes et al. (2012) have been mentioned as using novel animals and comparing BP, and, in the latter case, IS subjects to specific control conditions. These studies will provide the basis of the experimental design in the next chapter. Having seen that generic interpretations are easily accessible for singular generics subjects, the question arises as to whether more information about one particular animal could guide participants even more towards distinct representations of a novel animal's property, based on the wording with which they were introduced to it.

In the following chapter, I will present an experiment where the type of task that participants were asked to carry out is extended beyond a simple match-to-sample task. Instead, the test phase of the experiment employs a range of tests, targeting category-property links but also assumptions about essentialism of properties with regards to the animal category of which they are predicated.

[15] Data collection for these studies is currently under way.

*The age of **men** is over, the time of **the orc** has come.*

— Jackson *Gothmog, Lord of the Rings: The Return of the King*

5 Narratives about novel animal kinds

Contents

The previous two chapters have tested experimentally the role of morphosyntactic variation in generic subjects, first in a pseudoword paradigm that highlights the role of DEFINITENESS and NUMBER, and then in a series of match-to-sample studies, in which participants had access to world knowledge. The latter experiments especially have addressed genericity and kind reference with a focus on categorising and judging single instances of novel creatures on a case by case basis. This design was implemented to address the question of the influence of world knowledge and whether it helps participants in accessing more fine-grained distinctions in their representations of different generic subjects. The studies highlighted the importance of wording, not only in the test sentences but also in the prompts that introduce participants to the task. The results of this series of studies were also mixed and did not reliably support the hypothesised interactions. A further parallel was that these studies, like the *pseudoword* studies in Chapter 3, consistently showed a main effect of principally connected properties over statistically connected ones. Acknowledging that even though participants had access to world knowledge in the *kevtas* studies, they were still artificial.

To address this question, this chapter presents a study with a more narrative setting. The aim is to introduce novel animal kinds in a more realistic and less educational setting than in the previous studies (recall in particular concerns raised with regard to the intonation of the auditory cues). The studies rely on a well-established area of research in psychology that is interested in understanding concept acquisition from the standpoint of simulating naturalistic childhood development, based on both purely theoretical (e.g. Murphy & Medin, 1985) as well as experimental (e.g. Gelman & Markman, 1986) research. To further explore how morphosyntactic variations in generic subjects affects concept acquisition, this chapter will bring together arguments from the developmental literature and introduce a narrative approach.

5.1 A developmental approach

The main objective of this chapter is challenging the assumption that the most prominent generic subject, the bare plural (BP), is sufficient to provide relevant insight into how we generalise from instances to kinds, or whether we assume certain properties to be essential based on how we are linguistically introduced to the kind. Continuing this research from previous chapters, I will explore a developmental perspective on the previous experimental work.

Let us return to the idea of *Generics-as-Default*. This section will begin by looking at two distinct experimental approaches, one more firmly based in psychological approaches by Leslie et al. (2011) and another more firmly based in perspectives taken in linguistics by Lazaridou-Chatzigoga et al. (2015). I will address how their findings apply to research in developmental psychology by investigating how Gelman et al. (2010)'s studies fit into the research investigating *Generics-as-Default*. Methodologically, these two studies are distinct because they do not rely on measuring reaction times, which is typically used to argue in favour of *Generics-as-Default*, but instead use the relatively early age of acquisition of generics, compared with quantified and specific statements. As discussed in detail in Section 2.4, *Generics-as-Default* argues that generics are the cognitive default as opposed to specific or quantified statements (as proposed, adapted, and tested by Leslie et al., 2011; Hollander et al., 2009; Brandone et al., 2012, among others). For the purpose of understanding the implications for developmental research and particularly for the study presented in this chapter, let us briefly review its main characteristics here.

The underlying assumption that helped *Generics-as-Default* obtain a firm foothold in current developmental psychology is that any default cognitive process should require less effort compared to other types of cognitive processes. In the case of *Generics-as-Default*, this is exemplified by the automatic and easily accessible interpretation of statements as generic. Linguistically, this is underpinned by the absence of an overt marker of genericity, as opposed to overt pieces of morphology indicating a subset or a specific instance. Another argument that has been drawn upon in support of *Generics-as-Default* is the early age of acquisition of generic statements: Developmental research has shown that both quantified and specific statements are acquired later than generics. Finally, some studies have shown that quantified statements may be incorrectly recalled as generic, but generic statements are not incorrectly recalled as quantified (e.g. Sutherland et al., 2015; Leslie & Gelman, 2012). This is taken as further evidence that the fundamental form of storing world knowledge is in its generic form.

The first factor that I will focus on, which has been argued to be a strong piece of evidence for *Generics-as-Default*, is reaction time. Based on experimental results that show faster processing times and an earlier age of acquisition for generic statements compared to similar quantified expressions (e.g. *Barns are red* vs *Most barns are red*), psychologists argue that generics must be the default, and not specific or quantified statements. These faster reaction times are supposed to be indicative of the fact that generic statements do not require the conceptual system to resort to additional, and therefore more taxing, cognitive processes.

Within this context, recall that this interpretation of experimental data was met with some critical investigation from the linguistic community (in particular, Lazaridou-Chatzigoga & Stockall, 2013; Lazaridou-Chatzigoga et al., 2015). Seeing that linguists have tried to account for generics' variable truth-conditions and subset relations with modals or covert adverbs of quantification, such as the covert modal quantifier GEN, an attempt to reduce their complex semantics to faster processing times within a specific experimental design seems too simplistic. Linguists argue that with a more careful manipulation of the linguistic variables, researchers will arrive at more nuanced findings. These would require alternative explanations, which would be more closely aligned with the multi-faceted nature of the semantics of generics.

One such approach is presented in Lazaridou-Chatzigoga et al. (2015). Lazaridou-Chatzigoga & Stockall have addressed precisely the issues that were raised with regard to these previous experiments that claimed that *Generics-as-Default* is the most likely explanation for the special status of generics in the conceptual system and carefully redesigned the stimuli used in Leslie et al. (2011). To recall some background theory for these two studies, in Section 2.4 I discussed that *all* differs from the other two universal quantifiers *all the* and *every*. Lazaridou-Chatzigoga & Stockall (2013, p. 330f) argue that while all three quantifiers universally quantify the statement to which they apply, they possess underlying differences. This is noticeable when comparing the extent to which they allow for generic overgeneralisation (GOG). For instance, *all* has a corresponding generic interpretation, which is likely the reason this quantifier has been included as the universal quantifier in Leslie et al. (2011). Lazaridou-Chatzigoga & Stockall take this observation a step further and claim that this means that *all* then does not display GOG, but simply allows for its generic interpretation. By contrast, the other two universal quantifiers behave differently: Neither *all the* nor *every* can receive a generic interpretation (e.g. *All the ducks lay eggs*, *Every duck lays eggs*). Both of these quantifiers are either contextually restricted or interpreted with reference to a subkind.

In their studies, Lazaridou-Chatzigoga & Stockall aim to test whether the observed difference between generics and universally quantified statements, tested by comparing BPs with *all*, withstands further scrutiny. They introduced more variation in the subject form by also testing the two quantifiers, *all the* and *every*, in addition to the generic BP and the quantifier *all*, which were used in the original studies by Leslie et al.. To ensure that none of these quantifiers receive a generic interpretation, domain restriction was provided via a context within which the test sentences were presented. This more fine-grained test shows that instead of being able to claim that generic subjects are processed the fastest, their data are only able to support that a certain type of generic statement is processed faster than a certain type of universally quantified statement. Lazaridou-Chatzigoga & Stockall found that the supposed differences in processing speed do not hold for the wider range of wording they introduced in this replication. They argue that for the quantifiers *all the* and *every*, reaction times and judgements were dependent on the type of connection. It was only for "majority characteristic" and "minority characteristic" connections between category and property that they found the same results as Leslie et al.. The only significant interaction they found was between generic statements and *all*-quantified statements, and only when these sentences express majority characteristic properties.

In other words, they did not replicate Leslie et al. (2011)'s findings for *all*, and more generally for *Generics-as-Default*. In addition to these limitations, the findings in Leslie et al. (2011) are subject to further factors that affect the interpretation and judgements of generic vs universally quantified statements. Reading and reaction time are not solely determined by whether a statement is generic or quantified but instead, truth-value judgements are also influenced by the type of quantifier and the context within which they occur (Lazaridou-Chatzigoga et al., 2015, p. 488). In summary, Leslie's studies do not necessarily provide straight-forward support for *Generics-as-Default* but instead can also be interpreted under other accounts, e.g. domain restriction.

The focus of my thesis has been to explore the role of connection type between category and property in terms of which generic subject type they allow. However, these experimental findings regarding *Generics-as-Default* specifically provide an interesting parallel to the approach that I have previously taken. First of all, Lazaridou-Chatzigoga & Stockall expanded the variety of the quantifiers in Leslie's studies, rooted in the psychology tradition, while I have expanded the types of generic subjects. Second, the limitations that I have pointed out between generic subjects and connection type seem to be mirrored outside the realm of generic subjects. By including multiple universal quantifiers, Lazaridou-Chatzigoga & Stockall discover systematic limitations between quantifiers and the connection type they allow. This opens up the discussion for future research to consider the parallels between different types of generic subjects and quantified ones, and whether they can be extended to other subject types that come in various forms, e.g. specific subjects expressed via demonstrative determiners, which are also often used as a baseline for generic statements. This implies that there is a broader mechanism at play in which subtypes of specific morphosyntactic expressions of subjects are limited by the conceptual content expressed in the predicate of the sentence.

The conclusion then cannot be as straightforward as the idea that generics are processed more quickly than quantified statements. These experiments have shown that BP generics are processed faster than some types of universally quantified statements. While the argument in favour of *Generics-as-Default*, based on its quick processing times, is not verified, it is acknowledged that BPs render fast reaction times.

These tests are relevant beyond reactions times since the ease of processing is hypothesised to be related to its default status, and by extension its early acquisition. Returning to the questions that have driven the research for this thesis, much of the experimental design has relied on the differences **within** generic subjects. To further explore the role of morphosyntactic variations in generic subjects, while also considering arguments from the developmental *Generics-as-Default* literature, I will look at the role of generic language in novel concept acquisition under a developmental lens, as well as using a quantifier as a baseline instead of the previously used *this*.

A quick note on *this* as a non-generic baseline: The demonstrative *this* is often used as a baseline, referring to specific instances, against which the effect of generic language is compared. Likewise, the experimental set-up for the study in this chapter relies on the assumption that demonstratives connote individual reference. However, in certain contexts, they can also receive a generic interpretation. These "generic demonstratives" are felicitous in certain contexts, such as the following from Bowdle & Ward (1995, p. 33):

(1) [in front of a computer] This IBM ThinkPad is amazing!

Bowdle & Ward analyse these generic demonstratives and find that both singular and plural as well as proximal and distal demonstratives (*this, that, these, those*) are felicitous for presumed homogenous categories, i.e. often subordinate rather than basic kinds, that are familiar to the interlocutors. Note that plural distal demonstratives seem to be able to occur more freely (Bowdle & Ward, 1995, p. 32):

(2) A: My roommate just bought a Labrador.
B: Those labradors make great pets.

Related to both the question regarding *Generics-as-Default* from an age of acquisition standpoint, and using *this* as a baseline, consider Gelman et al. (2010), one of the formative studies for *Generics-as-Default* within developmental psychology. These experiments used a picture book that featured a novel animal (*zarpie*) and was designed to appeal to children. In linguistic terms, there were three versions of this picture book, corresponding to their three experimental conditions, differentiated by the wording of the subject form: The 'generic label' book introduced the novel kind in the BP as *zarpies*, the 'specific label' book introduced it as *this zarpie*, and the 'no label' condition provided only the demonstrative *this* to describe the novel animals throughout the whole book. Gelman et al. (2010) used a between participants design, such that no participant saw the new animal referred to in more than one way. This includes the title page of the book, which was called "All About Zarpies!" for the 'generic' condition. By contrast, the 'specific' picture book was titled "Look At This Zarpie!" and in parallel fashion, the 'no label' book featured the title "Look At This!".

This points to another way of categorising the books, namely by considering whether they introduced labels for the novel animal at all. This perspective enables us to introduce a different kind of hierarchy: Initially, the books are distinguished based on whether participants are given a label or not. If a label exists, then there are two types thereof to which a participant can be exposed, 'specific' or 'generic'. Compared to 'no labels', both specific and generic labels are expected to help categorise items, due to the labeling effect. By introducing a novel object with a label, as opposed to no label, children are better able to categorise instances of a new category. Plunkett et al. (2008) showed that the consistent use of a label, even for perceptually dissimilar objects, allows young children to group these objects together. This effect has been also observed in other studies and provides evidence that the mere presence of a label for a group of things enables people to draw inferences more quickly and confidently (e.g. Waxman & Markow, 1995; Gelman, 2004; Fulkerson & Waxman, 2007).

Bearing in mind the picture book methodology of Gelman et al. (2010), let us briefly return to the possible interpretations for *this*, and whether this would have affected their results. In light of independent research regarding the status of generic *this*, it seems unlikely that the generic interpretation should have been available in the Gelman et al. (2010) studies, and interfered with the hypotheses. As example (1) showed, singular proximal demonstratives require a specific context to be interpreted as generic. Further, within the context of carrying out a novel concept acquisition task, a demonstrative typically only receives specific reference.

As with the experiments in Chapters 3 and 4, the experiment presented in this chapter continues to challenge the assumption that the BP as the most prominent generic subject is sufficient to provide relevant insight into how we generalise from instances to kinds. Focusing on the minute morphosyntactic differences in how we are introduced to the kind helps us explore the unique expectations they set for the human conceptual system. Evidence from linguistic research, presented in Chapter 2 supports the assumption that the morphosyntactic form of the subject affects our understanding of a concept and its predicated properties. To make this point maximally salient, I have drawn on varying acceptability judgements for minimal pairs of generic statements. This rest of this section will briefly repeat relevant differences between the three types of generic subjects for count nouns in English as well as their interactions with connection types. These observations have informed the stimuli design for the *zarpies* study in this chapter, as they did for previous experiments in this thesis.

5.1.1 Variation in generic subjects by connection type

As discussed in Chapter 2, the form of the generic subject can be limited by the type of property that is expressed by the generic statement. In the case of principled connections (PCs) that express majority characteristics (3), all three generic subjects available for count nouns are acceptable (3-a)-(3-c). The only differences in judgement become apparent in the quantified version of the statement (3-d):

(3) a. Tigers are striped.
 b. A tiger is striped.
 c. The tiger is striped.
 d. Most/#All/?Some tigers are striped.

In this context, it bears repeating that people are prone to accepting universally quantified expressions even if not all critters possess the property. Part of the findings in Leslie et al. (2011) was that participants agree with statements such as *(All) ducks lay eggs*, even after being told that only healthy, mature, female ducks lay eggs, which picks out a minority subset of all ducks.

A similar pattern holds for definitional generic statements (4), where the distribution of acceptable quantifiers is the only indicator of the differences in subset relations. While the BP is the most commonly used form for the generic versions in (3), for (4), the definite singular (DS) subject in (4-c) is often judged to be better suited to convey the definitional aspect than the BP in (4-a), or the indefinite singular (IS) in (4-b):

(4) a. Triangles are three-sided.
b. A triangle is three-sided.
c. The triangle is three-sided.
d. *Most/All/*Some triangles are three-sided.

This preference becomes even more obvious when a PC is used with a kind-level predicate. To illustrate, consider the judgements in (5) (originally introduced as (45) in Section 2.3):

(5) a. ?Dodos are extinct.
b. *A dodo is extinct. [excluding subkind readings]
c. The dodo is extinct.
d. *Most/?All/*Some dodos are extinct.

While the BP in (5-a) is not unacceptable, it is the DS in (5-c) that receives the best judgement. This is crucial for the argument that the BP should not be the default generic subject and that differential judgements do not only mean that other subject forms apart from the BP are grammatical but rather, that other subject forms are preferred in combination with certain types of predicates. This holds for generic phrases with different types of kind-level predicates, which are generally preferred with a DS subject (e.g. *be extinct, be invented*).[1] (5-b) can be coerced into a subkind reading, as in *One kind of dodo is extinct*, but is not acceptable with the intended kind reference.

The picture of linguistic distributions that has emerged for PCs changes when considering statistical connections (SCs), as in (6). The preference for BP subjects, which is typically taken as a stand-in for generics in toto, seems more justified

[1] Another possible candidate for kind-level predicates is *be widespread*, but see Mueller-Reichau (2013) for further discussion.

for this type of category-property links. For majority statistical connections, the BP is the only acceptable generic subject:

(6) a. Barns are red.
b. *A barn is red.
c. *The barn is red.
d. Most/*All/Some barns are red.

There are other generalisations that rely on minority statistical connections, which are usually denoted as "striking property generics" (see e.g. Khemlani et al., 2007; Prasada & Dillingham, 2006, 2009; Prasada et al., 2013), as in (7). The distribution with regard to generic subjects is the same as for majority statistical connections. The difference appears in the quantified statements (7-d), expressing our world knowledge that only some mosquitoes carry the West Nile virus (in fact, less than 1%):

(7) a. Mosquitoes carry the West Nile virus.
b. *A mosquito carries the West Nile virus.
c. *The mosquito carries the West Nile virus.
d. *Most/*All/Some mosquitoes carry the West Nile virus.

While (6) and (7) seem to present a homogeneous picture for the distribution of generic subject types expressing generalisations of statistical properties, there are exceptions. Example (8) shows that properties that are present in the majority of the instances of a kind, but are affected by environmental factors, allow for an IS subject, but neither the BP nor the DS. Environmental factors are accidental, i.e. not principally connected to the kind. In this case, it means that if mole crabs were to live in a different type of suitable habitat, they would not have muddy feet. These seem to be the only concept-property connection that allows for this distribution of generic subjects (I thank Erin Saupe (p.c.) for these examples):[2]

[2]These judgements are controversial for some speakers. An alternative example is **Crocodiles are wet./{A/*The} crocodile is wet* or *A trout has wet scales*, which some speakers judge to be better. These properties are true of most, if not all instances of the kind but only because of external circumstances. Overall, the role of environmental factors to a category's properties is not

(8) a. ??Mole crabs have muddy feet.

b. A mole crab has muddy feet.

c. *The mole crab has muddy feet.

d. Most/*All/?Some mole crabs have muddy feet.

These examples together illustrate that PCs, as in (3) and (4), allow all three types of generic subjects, but merely statistically connected properties do not (6). Crucially, the BP may be the subject form that is most commonly used in generic phrases in general, but it is in fact not always the preferred form. This point is particularly clear in (5) and (8).

Finally, Lawler (1973) argues that singular generics must express necessary, inherent, or essential properties, as well as Burton-Roberts (1976)'s argument that singular generics need to carry normative force. This is illustrated in Lawler's well-known example, contrasting inherent (9) and accidental properties (10) and their acceptability with a singular subject:

(9) a. Madrigals are polyphonic.

b. A madrigal is polyphonic.

(10) a. Madrigals are popular.

b. *A madrigal is popular.

In summary, while quantifiers allow us to define specific subset relations with regard to the relative amount of instances of the kind that possess the property in question, generic subjects do not overtly mark the associated prevalence of their properties. However, there are hints in the acceptability distributions as to the type of connection that exists between kind and property. The reasons for these distributions relate back to assumptions about the underlying semantic representations (see Sections 2.2 and 2.3). Linguists have long been concerned with precisely these fine-grained differences, which are attributable to NUMBER

limited to animal kinds. Note that similar judgements are possible for *Leaves fall/{A/*The} leaf falls in the fall.* Here, the BP subject *Leaves* is grammatical, whereas the BP subject *Mole crabs* is not, or only in certain contexts or coerced readings.

and DEFINITENESS and we have good reason to assume that their underlying representations are different in terms of lattice-structures, if NUMBER is present (e.g. Link, 1983). This can, in turn, be related to their different covert projections in the subject DP.

Since the acquisition literature has focused almost exclusively on BP generic subjects, it is unclear whether early input supports the hypothesis that these distinct semantic representations raise distinct conceptual expectations. To understand the development of generalisations in the conceptual system fully, it is crucial to have a better understanding of the type of early linguistic input children are exposed to. While knowing the features of child-directed speech is one measure for investigating conceptual development, innate linguistic, and more specifically grammatical knowledge, may guide any type of input during conceptual development. The puzzle of induction persists in this analysis. Knowledge of the type of early input cannot entirely answer the question of acquiring a varied understanding of different types of generic subjects, but it can still play a role in determining how and when these distinctions can be learned. This is pertinent in light of the developmental findings of the *kevtas* studies in Chapter 4, which have shown the possibility of a comparatively late acquisition of SINGULAR generic subjects, and the DS in particular. In terms of the cues the conceptual system uses, experimental results in Chapter 3 have hinted at the possiblility that we rely on linguistic differences, at least in the absence of other information such as visual and/or auditory cues. The addition of this type of information, as illustrated by the *kevtas* studies, modulates the extent to which the conceptual system relies on linguistic variation.

5.1.2 Comparing generic subjects to quantifiers

The motivation for the *zarpies* study in this chapter is two-fold: Firstly, the study combines the comparison of multiple generic subjects, as is the basis of all studies in this thesis, with the majority quantifier *most*. The quantifier acts as a control condition and provides a different kind of baseline, similar to the universally-quantified sentences in both Leslie et al. (2011) and Lazaridou-Chatzigoga et al.

(2015). This sets this study apart from the other ones in this thesis, which have either compared the generic subjects against each other (Chapter 3) or against the proximal demonstrative *this*. Changing the baseline from *this* to *most* provides a better estimate of a baseline for properties that participants expect to be present in the majority of instances of a kind. In the *kevtas* studies, the specific baseline showed that all generic subjects were rated higher than it. Given that this research investigates differences within generic subjects, a preferable baseline is one that lies between a subject that support merely statistically connected properties (which are expected to be present in the majority of instances) and subjects that exclusively support principally connected properties (which are expected to be subject to less variation and should display tighter category-property links and higher essentialism). In other words, *most* is a more conservative baseline than *this*, since it behaves more in line with the critical conditions, in this case the generic subjects. *This*, on the other hand, is less conservative since it renders much lower, or much more different, results than expected from the generic subjects. Notably, Cimpian et al. (2010) have also used the quantifier *most* when comparing generic and non-generic language in studies using novel animals. They targeted people's understanding of animal-category links with implied prevalence estimation tasks as well as truth conditions tasks. Introducing participants to novel animals with either generic statements such as *Morseths have silver fur* or majority quantified statements, such as *Most morseths have silver fur*, is exactly parallel to the distinctions in wording that participants of the study presented in this chapter will see.

Secondly, it builds on the idea of investigating genericity from a developmental perspective. By using experimental material that is focused around only one novel animal and using a narrative technique that simulates picture books young children (and adults) encounter outside of a laboratory setting, it is more closely matched to an every-day setting of acquiring new conceptual knowledge than the single-exposure method of the *kevtas* studies. Additionally, due to the picture book style of the study, introducing multiple characteristics of just one animal kind, it resembles a reverse version of the training paradigms in Chapter 3: In those studies, multiple animals had one of two properties, and these properties were repeatedly mentioned in the training and testing phase. For the study in this chapter, it is not the property but the animal that is repeatedly introduced, with a different property each time.

Finally, while a comparison of the various generic subjects with a specific subject addresses the issues of *Generics-as-Default*, it does not do so by measuring reaction times. Reaction times have often been relied upon as one of the key indicators of default generalisations, yet the discussion above has noted the relevance of other features related to *Generics-as-Default*. Since this is a replication of an experiment whose results have been used to support *Generics-as-Default*, although in a slightly more indirect manner than other studies, the next section begins by focusing on whether the results are replicable with a wider range of generic subjects. This is similar to the approach taken by Lazaridou-Chatzigoga & Stockall when extending Leslie et al. (2011). If the results are comparable, then future research should incorporate means to measure reaction times.[3] For the initial replication, the main interest will be whether there are differences in the strengths of category-property links and the scores of essentialism of a property for a kind.

The rest of this chapter presents an extension of the experiments conducted by Gelman et al. (2010), using a set of stimuli that is linguistically more extensive (Section 5.2). This is based on the assumption that in the final state conceptual system, all generic subjects have distinct but equally accessible mental representations. Section 5.2.3 presents Experiment 13 (narrative), which follows up on novel findings with an experimental paradigm that targets concept acquisition differently and offers another perspective on the combination of learning about novel animals and the role of morphosyntax therein. Experiment 13 (narrative) uses measures of category-property links as well as essentialism to assess how the properties predicated of the *zarpies* have been conceptualised and whether participants judge them to be generalisable to other *zarpies*. The data are analysed both in the same fashion in

[3] That being said, reaction times have been collected and analysed for Experiments 1-5 (DEF-NUM) and the follow up studies in Chapter 6. I will return to observations related to reaction times in the next chapter.

Gelman et al. (2010) and also through a post-hoc item analysis, looking at wording-independent response patterns, which did not form part of the initial hypothesis (Section 5.2.4). Uncovering additional patterns shifted the analysis to focus on an alternative explanation, as opposed to the hypothesised explanation that solely focuses on the morphosyntactic differences of generic subjects. Instead, the role of similarities of properties in basic-level kinds, which are inherited from superordinate kinds, is highlighted (Section 5.3). The implications for future research and a specific experiment, to be presented in Chapter 6, are discussed in Section 5.4.

5.2 Encoding properties of novel animals: All About Zarpies

In light of the experimental results reviewed in Section 5.1, this section will address in more specific terms the limitations regarding the stimuli used in previous psychology experiments targeting generic language. I will discuss the missing observations due to a 'generic label' condition where generics are only represented by the prominently used BP. One promising way this chapter builds on the knowledge about the effect of different generic subjects is by extending previous experiments that already tap into this question.

As briefly introduced in Section 5.1, Gelman et al. (2010)'s *zarpies* study consists of a set of three connected experiments, with two types of participants, four-year-olds and adults. This section will introduce the first of these studies, which recruited adult participants. These experiments gave the strongest results and are thus an ideal starting point to look at more fine-grained linguistic distinctions.[4]

[4]It might be interesting to consider briefly the design of the images for the *zarpies* in the picture book. Gelman et al. discusses that the images were constructed around a prototype. The participants do not see the prototype itself, but are expected to form a category structure around it based on the variations that they are exposed to. Looking at the early prototype theory literature, and in an anticipatory link to the discussion in Section 6.2, bear in mind that Rosch & Mervis (1975, p. 595ff) argue, based on their Experiment 5, that in artificial categories "items that have greater degree of family resemblance with the members of their own set are learned more rapidly, identified more rapidly even after practice, and judged as more prototypical members of the category". Artificial categories are partly similar to this experiment as the *zarpies* are a novel category. They differ, though, in that *zarpies* are an animal kind, and therefore, participants will have preconceptions about the animal's underlying structure. Thus, they lay the baseline assumption that any form of resemblance, be it visual or a string of letters and digits, raises

After being introduced to the *zarpies* via reading the picture book, the test phase of the study consisted of three subsequent components: a category-property component, an essentialism composite, and a memory task. Both Gelman et al. and the extension presented in this chapter exploit the fact that generic language is not only a reflection of our understanding of the world but, in fact, performs an educational function by teaching us about the relationship between objects and their properties. Rather than passively accepting generalisations, we have expectations as to the kind of connection, be it principled, statistical or even striking (Gelman et al., 2010; Cimpian & Markman, 2009; Fuellenbach, 2017), implied by each linguistic form.

Moving beyond a conflated 'generic label', I will extend Gelman et al. (2010) to a four-way subject design by adding the two singular generic subjects IS and DS, and a *most*-quantified control condition (MQ). These experiments can thereby highlight the role of variation in morphosyntax by changing the original subject forms from 'no-label', 'specific-label' and 'generic label' to 'BP', 'IS', 'DS', and 'MQ'.

The remaining set-up was unchanged. In a between-participants wording manipulation of the picture book, participants learned different properties about the novel animal *zarpie*. The only variable that was manipulated in all of the studies is the wording of the subject, leading to distinct subject types, e.g. *{Zarpies, This zarpie, This} hide(s) behind fences* or *{Zarpies, This zarpie, This} hate(s) ice-cream*. They were then asked questions targeting their understanding of the new animal kind from the same test booklet as in the original studies. Its three components are detailed below.

5.2.1 Testing category-property links and essentialism

In the study's test phase, participants were presented with a tripartite booklet, consisting of three components targeting (1) category-property links, (2) beliefs about the essentialism of the predicated properties, and (3) a memory task to ensure they recalled the information conveyed via the stimuli throughout the experiment.

expectations as to the internal organisation of a novel category. Further, these expectations are modulated by overlap between instances. Therefore, we can expect that participants will form rich internal category structures of the *zarpies*-instances, both based on the visual input and the predicated properties that are explicitly attributed to them.

Table 5.1: Questions for participants in Gelman et al. (2010)'s test booklet, grouped by test battery to which they are attributed.
Adapted from Gelman et al. (2010, p. 279), their Table 3

TASK	# ITEMS	EXAMPLE
Category-property composite		
Categorization	4	"Can you find another one?"
Familiar induction	6	"Do you think this zarpie *P*-s"
Explanation (scope)	4	"Why is this *P*-ing?"
Essentialism composite		
Explanation (content)	4	"Why is this *P*-ing?"
Stability	6	*True at 4 years old? True always?*
Inheritance	4	*Birth mother vs adoptive mother*
Novel induction	6	*see 'familiar induction'*
Match taxonomic/thematic/shape	4	*Match with one of three choices*
Memory composite		
Picture memory	4	*"Which of these two zarpies was in the book?"*
No label (subjects)	17	*Correct subject label*
No label (predicates)	58	*Correct predicate label*

The full set of tasks is presented in Table 5.1, grouped by the composite to which they belonged. All questions except the two 'no label' ones in the memory composite were included in the replication (Experiment 13 (narrative)). The test booklet presented to participants is attached in Appendix D.4.

Let us consider each of the question batteries in turn. First, the category-property component measured "the extent to which [participants] linked the category to the properties expressed" (Gelman et al., 2010, p. 273) and is crucial for this version of the study since it directly targets the relationship between the generic subject and its property. With it testing the strength between subject and property, I hypothesise, in line with Gelman et al., that when a connection between category and property is judged as strong, this "impl[ies] that a range of zarpies (zarpies in general) have this property. Thus (...) [participants] did not restrict this property to the instance at hand, but instead readily generalized this fact to other zarpies" (2010, p. 284). Thereby, BP generic subjects licensed high expectations in the face of low prevalence. This effect of generic language has been observed independently by Rhodes et al. (2012) who ran a replication of the original *zarpies* study including both BP and IS generic subjects. These studies compared BPs to the specific baseline *this* (with child participants), and compared only the IS generic (but not the BP) to the specific baseline *this* (with adult participants). Rhodes et al. (2012) found that the child participants treated IS generics like BP generics and both child and adult participants treated IS generics different than the baseline *this*.

Second, the essentialism composite directly feeds into these questions by assessing "the extent to which [the participants] treated the category as constituting an essentialized kind" (2010, p. 273). It raises the issue of whether participants believe in a shared, underlying commonality among members of a kind. These features are "a cluster of non-obvious, inherent properties" (2010, p. 274), at least for natural kinds.[5] These essentialist beliefs are proposed to be stronger if the subject is presented in generic form, as this implies that the category has properties with an underlying basis that connects the property to the members (and thereby instantiations) of the kind. Compared to the specific-label and the no-label condition in Gelman et al., participants "treat the category as more stable, more inductively rich, etc." (2010, p. 285) with a generic label. If generic subjects give us access to understanding a category in general, as opposed to providing data points about an observation of a specific instance, this link between category and property is expected to be stronger if the information is expressed in a linguistic form that matches the generalisation. In other words, generic statements, by virtue of expressing generalisations, are hypothesised to imply tighter category-property links than linguistic forms that merely support episodic observations about individuals. I expect to find graded response patterns based on the type of generic subject with which participants were introduced to the category. Crucially, I do not expect that the BP is representative of all properties that can be expressed by the morphosyntactic forms available in English. Instead, I argue that they carry distinct semantic assumptions. In other words, I am not solely interested in the strength of the link but also in

[5]At this point, I leave aside the question of how these properties would be ascribed to underspecified categories. This concerns, for instance, the properties that would be introduced in a pseudoword paradigm, as in Chapter 3, or for other types of categories, such as artefacts.

expectations as to the kind of link, which can be principled, accidental or causal (Prasada & Dillingham, 2006, 2009).

Finally, the memory component serves to test whether participants were mindful of the wording and information provided, thereby ensuring that the observed effects can truly be attributed to the wording of the subjects. Participants are asked to recall the information that was provided in the picture book and fill it in, as precisely as possible, in a blank version of the picture book.

One notable characteristic is that these studies exclusively use animal kinds for novel concept acquisition, unlike some of the other novel kind studies run previously (e.g. Fuellenbach, 2017, which specifically makes an effort to underspecify the type of novel category). When exploring the link between generic language and essentialism, which was the goal of these previous studies (e.g. Gelman et al., 2010; Hollander et al., 2009; Gelman & Tardif, 1998), this is a useful consideration. Not only do animal categories show higher essentialism, but also generic language is more prominently used to talk about these kinds of categories (Gelman et al., 1998; Hollander et al., 2009).

At this point, it should be reiterated that different kinds of categories (e.g. natural kinds, social kinds, artifacts etc.) behave differently in generalising statements. Experiment 13 (narrative) will follow Gelman et al. (2010) and focus on just natural kinds. In addition to that, the interpretation of the property depends partially on the kind of category of which it is predicated (e.g. colour terms predicated of animal kinds are more likely to be expected to be principally connected than colour terms predicated of artefacts). In Experiment 13 (narrative), I continue to restrict the stimuli to novel animal creatures and thereby restrict the interpretation of how certain properties usually relate to them. Consequently, I expect participants to extrapolate from their current knowledge on the available types of connections between animals and the known properties, even without any further information about how the (known) properties are in fact related to the novel animal kind.

While I emphasise heavily that different generic subjects bring with them distinct expectations, note that Gelman et al. have also included in their discussion that the

BP is not the only generic subject and explicitly mention the linguistic differences. They differentiate between the BP and IS, saying that "[t]he bare plural form of the generic can be used to express either principled connections (e.g. 'dogs are four-legged' because four-leggedness is a natural state for dogs) or statistical connections (e.g. 'barns are red' because the majority of barns happen to be red), whereas the indefinite singular more clearly implicates principled connections" (Gelman et al., 2010, p. 296, emphasis added).[6] Based on their observations, one can hypothesise an interaction between wording and connection type, leading to differentiated effects for each combination of subject and property type. This is exactly the kind of graded strength effect hypothesised for this replication, with the difference that the DS is also included to account for the full paradigm of generic subjects for count nouns in English.

My research takes the finding that people tend to generalise to other instances of a kind when presented with generic language, and further asks what types of expectations they raise when presented with different types of generic language. In other words: Based on our assumptions about the linguistic properties of IS or DS generic subjects, do participants still expect the same kinds of connections between the subject and its property or do we find more fine-grained distinctions?

5.2.2 Predictions based on the subject's NUMBER

In terms of the graded strength effects mentioned above, I predict that the strength of the connection between subject and property will be modulated based on the form of the subject with which the *zarpie* is introduced in the picture book. In the first instance, the results for the BP generic should reflect the data in Gelman et al. (2010). They should show tight category-property links and high essentialist beliefs and can thereby offer a baseline for the singular generic conditions, i.e. a generic against which to compare the other generic subjects. This would make the comparison similar to the pseudowords studies.

[6] Note that the DS is not mentioned here.

For the singular generic subjects, the predictions rely both on applying the arguments made in Sections 2.2 and 2.3. I will revisit these for each subject type and relate them to the predictions for Experiment 13 (narrative).

Of the singular generic subjects in Experiments 1-5 (DEF-NUM), some studies showed preliminary evidence, in line with the hypotheses based on the literature review, that IS subjects can lead participants to have higher expectations of the connection between subject and property to be principled than when the properties occurred with a BP subject. Because IS subjects refer to individuals rather than groups of individuals, participants should expect that a property introduced with an IS subject be principally connected to new objects since the semantics of the IS required that each individual possess the property.[7] Even though Experiment 13 (narrative) does not explicitly test principled and statistical connections, the semantic representation of the IS is still such that each instance of the kind is expected to possess the predicated property. Therefore, in the questions raised in the test booklet, IS subjects should show the strongest connections between the subject and property. Specifically, this means that the IS is hypothesised to score higher in the category-property links composite and the essentialism composite than the BP.

In the pseudoword experiments, the DS contrasted with the IS and by comparing some of the studies, DS subjects displayed an asymmetrical pattern. Properties that are introduced with a DS subject should lead participants to have lower expectations of the connection to be statistical. Following the hypothesis that DS subjects directly refer to kinds, they are linguistically distinct from the IS and BP in that it is argued that their NUMBER value is neither SINGULAR nor PLURAL, but numberless. Syntactically, this means that while both IS and BP have a NUMBER-phrase, #P, in the structure of their subject DP, this covert projection is lacking for kind-referring DPs (Borik & Espinal, 2012, 2015, see also Borer,

[7] These assumptions have been laid out in more theoretical detail in various publications, e.g. Burton-Roberts, 1976.

2005).[8] Its surface form is identical to the definite singular, i.e. the underlying structure of a subject which is made up of a definite determiner and a noun, with episodic reference. The kind-referring equivalent is therefore often assumed to also be SINGULAR. I argued that in the studies in Chapter 3 that indicated patterns in line with the hypotheses, participants came to expect that a property introduced with a DS subject was less likely to be statistically connected to a new object. This is because these types of generalisations about kinds relay information about their essential or causal properties.

In this paradigm, DS subjects should therefore support the belief that the connection is less likely to be statistical, since properties predicated of kinds, and not instances thereof, do not rely on statistical connections. In terms of variables that are measured in the test booklet, any property that could merely just happen to be predicated of the *zarpies* in the book, should receive lower scores in both the category-property links and the essentialism composite. If the DS is interpreted in its kind-referring function, the DS will score high in the category-property composite, indicating tight links between the category and its properties, as well as higher scores in the essentialism composite than the BP. However, a DS subject might be more difficult to interpret generically than other subject types. Due to the high level of ambiguity between generic and episodic interpretations, participants might also give DS subjects much lower scores, in line with a specific interpretation of 'the zarpie'. This experiment tests whether this assumption holds true experimentally, i.e. by assessing how accessible a generic interpretation is in a picture book as well as the exception tolerance of generics.[9]

As this thesis teases apart effects due to statistical knowledge as opposed to principled connections, this extension features a *most*-quantified condition, which

[8] While we may reanalyse a kind-referring definite 'singular' subject as a 'definite numberless' (DN), this theoretical distinction has no direct bearing on the argument presented here. Thus, we continue to refer to it as a DS subject. Where the difference between generic and non-generic definite singular/numberless subjects is crucial, this distinction should be made explicit.

[9] Note further that with the correct intonation, the DS subject very much sounds like instructional, child- directed speech, e.g. "The elephant $_{\text{pause}}$ lives in Africa". For a more detailed overview on the relationship between semantic interpretation and pitch, see also Kratzer & Selkirk (2007) and Longobardi (2000, 694). Further discussion on the interpretations of DS subjects that are phonologically accented can be found in Allerton & Cruttenden (1979).

specifically provides a measure to assess where majority statistical properties lie in comparison to PCs. PCs are expected to be equally high as *most*. This is due to generic overgeneralisation (GOG, as introduced in Section 2.4), where *all*-quantified properties are expected to be prevalent upwards of 90% of the instances (Leslie et al. (2011), but see also further research, such as Sutherland et al. (2015)). This condition should provide a baseline against which to compare the three generic connections, as it is argued to only tap into statistical connections with majority prevalent properties. In terms of experimental predictions, subjects that are quantified with a majority quantifier should behave similarly to the BP in that the overall scores should be high because the predicated properties are expected to be prevalent in most instances. However, this should not necessarily lead to the assumption that they are essential to the category. Scores in the essentialism composite should be lower than for IS and DS subjects, but similar to the BP, which also allows for merely statistically connected properties.

In summary, this experiment seeks to provide a linguistic perspective on the well-known puzzle of induction, in acquisition and communication, and to ultimately highlight the role of linguistics in psychological experiments. In other words: can linguistic cues enable us to better understand the links between categories and their properties? The subject condition was fixed within each picture book, rendering a between-subject comparison. This design allows for more detailed comparisons between generic subjects, and is based on and extends Rhodes et al. (2012). By including questions that target category-property links and beliefs about essentialism, the results can help to uncover subtle semantic differences in the subject form by the participants' distinct behavioural responses.

5.2.3 Experiment 13 (narrative): Adult participants

5.2.3.1 Participants

Fifty-six adults (33 female, 23 male) participated in the experiment. They were assigned one of four conditions (n=14 each). The adults were mostly undergraduate or graduate students at the University of Oxford, gave informed consent to participation, and received either course credit (3 credits) or financial compensation (£7.50) for their time.

5.2.3.2 Items

The items were adapted from Gelman et al. (2010)'s study. I retained the condition with BP subjects and included two new generic subjects and a quantified subject condition. The way that the three generic subjects contrast and interact with the different types of properties was the focus of this extension of the original study. The *most*-quantified subject was included to act as a baseline for statistically connected properties. The original studies included the specific condition *This zarpie...* as a baseline condition. The change from *this* to *most* was due to my interest in comparing the BP, which allows for statistically connected properties, with a quantifier, i.e. a linguistic marker that explicitly allows for only statistically connected properties, in this case a majority subset. This decision was further based on the specific comparison of differences between generic subjects, while a specific baseline condition is expected to be less generalisable than each of the three generic subjects.[10]

This replication study kept the between-subjects design. As in the original study, there were 17 zarpies in the book about which participants learned properties and three zarpies on the cover that were not repeated in the picture book. The sentences were formulaic, i.e. they did not vary throughout the book, and followed the form:
Look at this zarpie!
{Zarpies/A zarpie/The zarpie/Most zarpies} hide(s) behind fences.
This text was presented on one page, such that the BP generic prompt for this example looked as in Figure 5.1, followed by a picture illustrating a *zarpie* and the aforementioned property on the next page, in this case as presented in Figure 5.2.[11]

[10] To address any discrepancies, the results from the original study's baseline condition have been plotted alongside the findings from Experiment 13 (narrative). Future studies should also include the original baseline condition *this* to allow for comparisons between the studies and the different generic subject forms.

[11] Relatedly, I would be interested in the different types of connections these properties could have with the kind. To illustrate, 'has stripes at the bottom of their feet' is more likely to be a

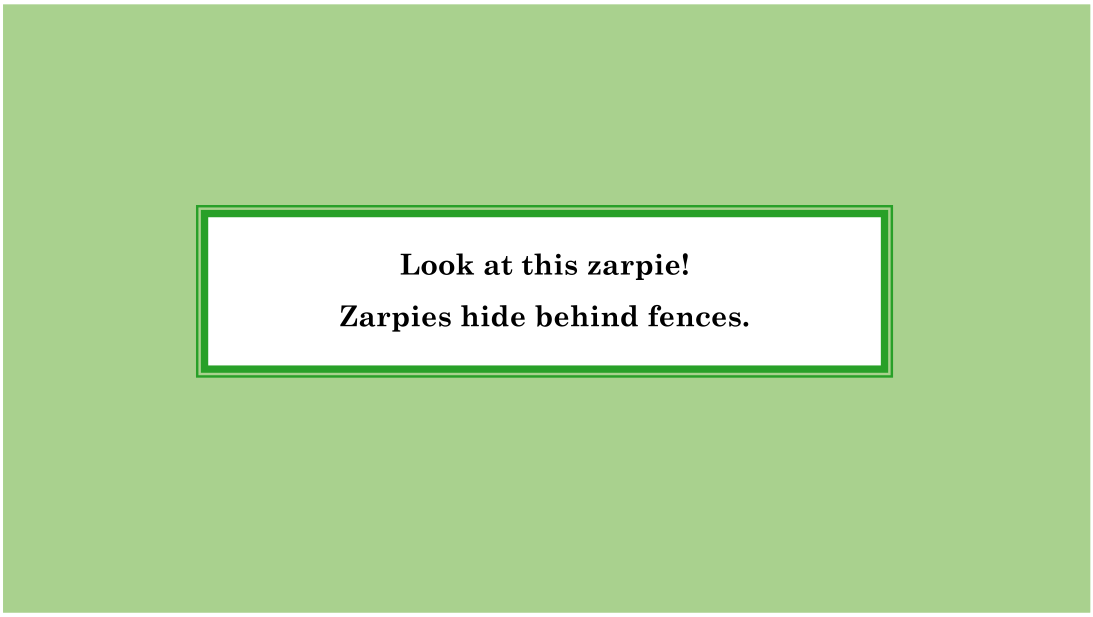


Figure 5.1: Sample prompt from Zarpies picture book, BP generic condition.

Figure 5.2: Sample image from Zarpies picture book, BP generic condition.

5.2.3.3 Procedure

Participants were tested in a quiet lab environment, either individually or in small groups of up to four people. Participants received the *Participant Information Sheet* (see all materials used in this study in full in Appendix D, and the PIS in Section D.1) via email and in person before the start of the experiment. They were then encouraged to ask any questions about the experimental procedure that they might have.

The experiment itself consisted of four parts, replicating the methods and procedures in Gelman et al. (2010)'s Study 1. First, participants read through the picture book for five minutes. They were instructed to continue reading through the book for the entirety of the five minutes, even if they finished reading it in less than the allocated time. Participants then moved on to a distractor task, solving as many multi-digit multiplication problems as possible during a set time period of four minutes.

The third part was the testing phrase during which participants were asked to fill out the test booklet (with the question introduced in Section 5.2.1 above and the full test booklet is attached in Appendix D.4). It targeted multiple aspects of the novel animal kind that was introduced during the reading of the picture book. To analyse these differences between the generic subjects predicted above, all three questions from the category-property composite remain in the test booklet. Similarly, participants were asked to respond to all five questions from the essentialism composite. In this part of the test battery, I expected the IS to score significantly higher than BP plurals. The 'explanation (scope)', 'stability', 'inheritance', and 'novel induction' tasks will provide insights that are similar to questions that we have previously asked to understand whether a connection between subject and property is principled or statistical. Third, the picture memory task from the memory composite was included to check whether participants have

---

principled connection (PC) for animal kinds, whereas 'hates ice-cream' is more likely to be an idiosyncratic property that is statistically prevalent. However, this raises the question of which connections can a priori be deemed as principled and as statistical as part of the data analysis. A clearer distinction between these properties will be considered in future experiments.

paid attention to the book in general, as well as to the zarpies' properties and the wording more specifically.

Finally, participants filled out an essentialism questionnaire that assessed their beliefs about how essential, innate or stable they judged certain properties to be for people in general. The questions for the essentialism questionnaire are based on Gelman et al. (2007, 2010), and Haslam et al. (2000). The version used in this replication was a reduced variant of the questionnaire used in the 2010 studies, including 80 of the 140 items, or four out of the seven 20-question categories, which were randomly chosen. Participants rated how much they agreed with statements on a 6-point Likert scale. Its purpose was to find a baseline of essentialism within and between groups to allow for adjusted comparison between the different conditions, if necessary. Note that it also allows us to make inferences about how language might affect alleged baseline beliefs, which I further address in Section 5.2.3.5 below.[12] The entire experiment lasted between 30-45 minutes.

5.2.3.4 Analysis and coding

Coding was carried out in line with the original studies to ascertain that the original results were replicated. The originally labeled 'generic'-condition was renamed 'BP-generic', and the two novel generic subject conditions are presented here as 'IS-generic' and 'DS-generic'. The *most*-quantified condition is labeled 'MQ'.

Participant responses were analysed based on whether they generalised a property to a novel instance or chose an answer that judges a property as essential to the animal a score. Those answers were codes as "1". All other responses were given a score of "0". To illustrate, provided the information that *zarpies* sing, if a participant answered that a *zarpie* that was adopted by dogs, which bark, sings, the property of 'singing' was extended to another instance, thus receiving a score of "1". If a participant responded to the same question by answering that this

[12] In discussions with Susan Gelman later on, it became clear that they also found 140 items to be too many and reduced the essentialism questionnaire even further. This shorter version was used in the *kevtas* experiments with child participants in Section 4.4.3. While the essentialism questionnaire used for the *zarpies* study is shorter than essentialism questionnaires used in the above-mentioned studies, it is still larger than the one for the *kevtas* studies. See this version in full in Appendix D.5.

zarpie barks, their answer received a score of "0". For a full breakdown of the questions that targeted category-property links and those that targeted essentialism, confer Table 5.1 above. The participant's scores were converted into percentage, both separately for responses for the questions targeting category-property links and questions targeting essentialism, as well as jointly for an overall score. The 'explanation scope' and 'explanation content' sections were coded by both the main experimenter and a second coder. Intercoder agreement was 92.86% for the 'explanation scope' and 85.71% for the 'explanation content' section and any discrepancies were resolved by discussion.

Some additional analyses were carried out to test further hypotheses, complementing those in the original studies. This applies to the binary coding of the "match taxonomic/thematic/shape" questions in the original 2010 studies, which was extended in certain categories to catch the full range of possible answers. In the original studies, participants received one point for the 'taxonomic' explanation and zero points for either the 'shape' or the 'property' match response. In this analysis, alternative answers participants selected were taken into consideration, so as to avoid scoring '0' when a participant chose either the 'shape' or the associated 'property' match, instead of the 'taxonomic' option. Instead, both alternative answers were given independent scores. This means that for each wording condition, there were independent scores for the 'taxonomic', the 'shape', and the 'property' match. Further, for the "match taxonomic/thematic/shape" task, it was not possible to know why participants chose one option over the other two. Thus, special attention was paid to the justifications in the 'explanation' tasks to see if there were patterns that emerge for specific items or participants.[13]

[13]Moreover, thinking about how this line of research connects to the pseudoword studies in Chapter 3, let us once again consider connection types. While the original studies looked at how essential participants judged the properties to be, another way to look at the generic statements here is to consider the likelihood of a connection being principled or statistical. To illustrate, 'has stripes at the bottom of their feet' is more likely to be a PC for animal kinds, whereas 'hates ice-cream' is more likely to be an idiosyncratic property that is statistically prevalent. The question as to which connections can *a priori* be deemed as principled and as statistical arose as part of the data analysis. Having had a closer look into the expectations as to the connection type since data collection, it appears that only a subset of the stimuli originally categorised by Prasada & Dillinghamas principled or statistical do raise these expectations consistently and sufficiently distinct. For the *kevtas* studies in Chapter 4, only normed properties were included. This was

A oneway ANOVA was conducted to establish whether the differences in responses (by participant) for the questions in the picture book based on the 4 wording conditions (wording: BP vs IS vs DS vs 'most', between participants) were significant. There was a significant effect of wording condition of the different subject types on the scores in the test booklet ($F(3,52) = 9.7562$, $p < .0001$). These effects are illustrated in Figure 5.3.

Planned comparisons (Tukey HSD) indicated that the DS differed from all other conditions. Compared to DS subjects, participants were significantly more likely to generalise properties if they were trained on BP subjects ($p < .0001$), *most* ($p = .0004$), and IS subjects ($p = .0012$). No other wording conditions differed significantly (all p values between .8359 - .986).

Overall, the data from this study showed that participants selected responses that generalised properties to novel instances for the BP, IS, and MQ conditions. The responses for these three wording conditions also showed that participants chose options in the test booklet that assume that the predicated properties are essential to *zarpies*. All results will be discussed separately by responses given for questions targeting category-property links, essentialism, and memory recall below.

The results of the BP wording condition were similar to the results of the original Gelman et al. study. The responses for the DS wording condition differed the most, having the lowest scores for questions targeting whether a property was generalised to novel instances and questions targeting whether a property was seen as essential to the animal. A summary table of the scores for questions in the test booklet to which participants responded in line with the option that rates the zarpies properties as essential and where they chose the response option that extends the predicated property to novel instances, for all four wording conditions, is given in Table 5.2. Note in particular the large standard deviation and range for DS subjects.

initially assessed by testing whether a hypothesised PC would evoke normative force, as illustrated by the judgements in (i)-(ii) (see Prasada & Dillingham, 2006, 2009):

(i) Tigers should be striped.

(ii) *Barns should be red.

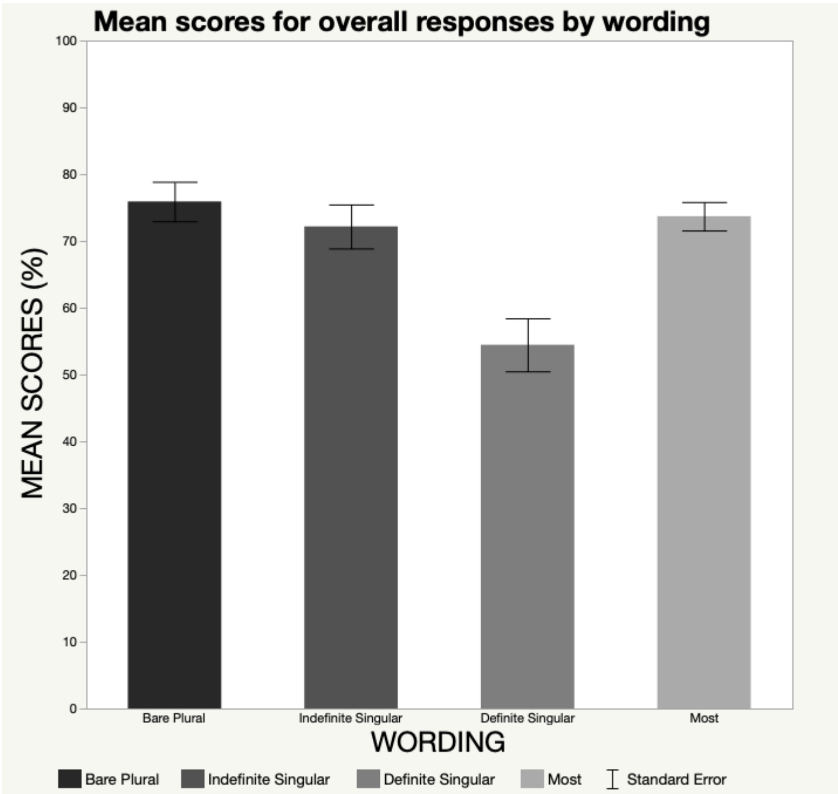


Figure 5.3: Mean scores of category-property links, essentialism, and memory questions for Experiment 13 (narrative), by wording condition. Error bars = ± 1 SE from the mean.

Table 5.2: Summary statistics of scores given in test booklet for Experiment 13 (narrative). The responses were given in response to questions targeting participants' understanding of properties as essential as well as generalising predicated properties to novel instances, by wording

	BP	**IS**	**DS**	**MQ**
Mean	72.62	69.90	53.06	68.88
Standard Error	3.47	3.00	3.94	2.68
Standard Deviation	12.97	11.23	14.75	10.04
Range	38.10	38.10	52.38	35.71
Minimum	50.00	47.62	28.57	52.38
Maximum	88.10	85.71	80.95	88.10

The scores for category-property links of Gelman et al. (2010) just outside one SE of the replication study. Participants were highly likely to expect properties associated with *zarpies* to extend from the instances presented in the picture book to novel instances in the test booklet. In order to analyse the results by their individual test batteries, the next section will analyse and discuss participant responses looking at each of the composites in turn, starting with the questions that targeted category-property links, then those that tested beliefs about essentialism, and finally the memory component.

5.2.3.5 Analysis of each test battery

In order to provide a detailed comparison of the current results with the original generic condition in Gelman et al.'s study, the results obtained in the BP condition were isolated for all three task types. The category-property links scored 87.24% (SE = 4.57; original study: around 82%), essentialism at 63.10% (SE = 4.43; original study: around 60%), and memory at 76.79% (SE = 5.54; original study: around 78%).[14] The results from the original study were mostly replicated, with the scores for category-property links in the 2010 study almost being within one SE of the replication study. More interesting for the current research question, however, is an analysis of how the different wording conditions compared to each other.

Inferences about category-property links This section provides an analysis of the responses given for questions of the section of the test booklet focusing specifically on category-property links. A oneway ANOVA was conducted to establish whether the differences in responses for the questions of the composite in the test booklet specifically targeting category-property links based on the 4 wording conditions (wording: BP vs IS vs DS vs 'most', between participants) were significant. This analysis indicated a significant effect of wording condition of the

[14] Gelman et al. note that across conditions, recall was over 75%, and the current responses are also about 75%. Most errors were related to recall of the predicate. The current findings, as will be explained in the **Memory** section below, show that in the DS subject condition, more errors occurred for the subject form compared to recall of the other subject conditions.

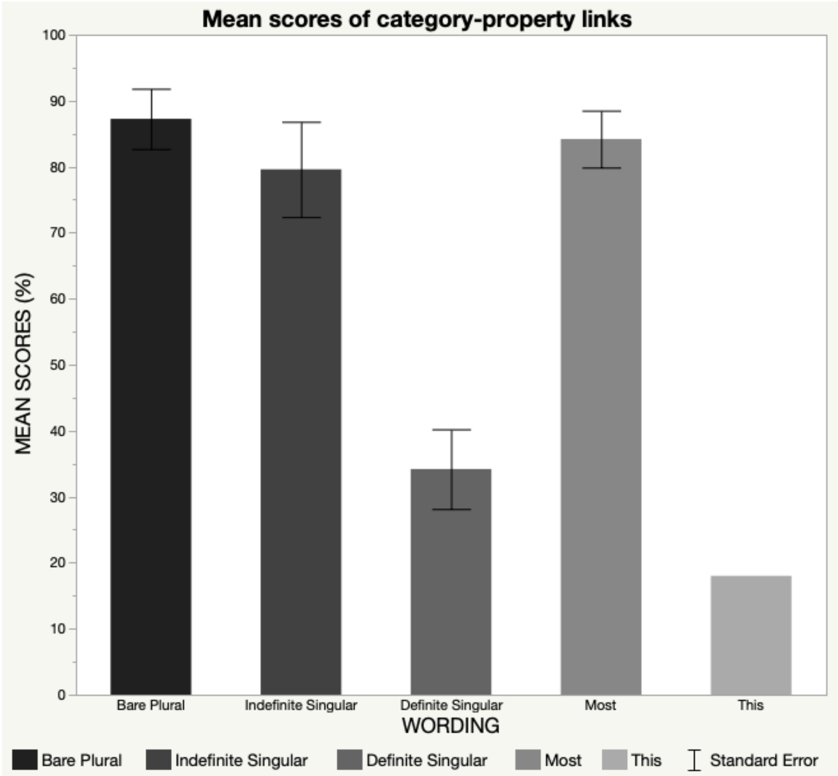


Figure 5.4: Mean scores of Category-Property Composite (Experiment 13 (narrative)), by wording. Error bars = ± 1 SE from the mean.

different subject types on the scores in the category-property links component of the test booklet (F(3,52) = 19.436, $p < .0001$), as plotted in Figure 5.4.

Planned comparisons (Tukey HSD) indicated that, compared to DS subjects, participants were significantly more likely to assume strong category-property links if they were trained on BP subjects, 'most', and IS subjects (all p values $< .0001$). There are no differences between BP, IS, and 'most'-quantified subjects (all p values between .7745 - .9807).

Note that the responses for questions targeting category-property links were similar to Gelman et al.'s results when the *zarpies* were introduced with a BP subject. Equally, when the *zarpies* were introduced with an IS subject, participants expected the properties ascribed to them to also hold true of novel instances, i.e. participants were likely to generalise and extend them to previously unseen instances.

I predicted that the IS subject would lead to higher scores in the category-property composite of the test booklet (as defined in Table 5.1 above), indicative of the assumption that IS generics raise expectations of principled connections between kind and property. However, the scores for IS subjects were not significantly higher, compared to the BP ($p = .7745$).

Seeing that the DS subjects scored significantly lower in the category-property composite than the other two generic subjects, and is overall significantly different from all other subject conditions ($p < .0001$), *the zarpie* seems to be mostly interpreted non-generically. This might be due to participants understanding the picture book as being written in a "narrative present" reading. Whether this is a valid alternative could be investigated by looking further into interpretations of other children's picture books, similar to the corpus analysis in Chapter 2.8. Recall that the picture books that were analysed, as well as the corpus data, feature very few cases of DS subjects. Yet, within the context of generic BP subjects (and pronouns with generic reference), these DS cases might be more easily interpreted as generic, while the *zarpies* picture books in this study only featured one subject type.

Further analyses show that the DS subjects show a wide-ranging distribution (scoring between 12-34 out of 42 possible points, mean 22.3, median 23.5), such that some participants rated the category-property links significantly lower than participants in the BP and IS subject conditions. This pushed the overall acceptance rate of category property links down compared to all other conditions. Other participants rated them equally high as the BP and IS subjects. I propose that the DS allows for kind reference, as theoretically stipulated, but that it is more ambiguous between the individual and kind-reading than the other subject types in this context.

The MQ subject condition was included as a baseline comparison for statistical measures, since participants should extend the existence of properties of one instance of a zarpie to others when being told that *most zarpies* have said property. The category-property link scores are similar to the BP and the IS, which is hypothesised to allow for averaging over groups and statistical connections. Therefore, their similar behaviour supports the hypothesis that majority statistical properties can be expressed with BP subjects, similar to the overt quantifier *most*. By contrast, the DS scores lowest in the category-property links (35.71%, see Figure 5.3), and is significantly different from all other subject conditions ($p < .0001$).

In Figure 5.4, the data from the control condition 'this' from the original Gelman et al. (2010) studies are included for comparative purposes (however no error bars are given as I did not have access to the full data set. See the original paper for a full analysis). If the strength of the link between category and property, as assessed by the questions in the test booklet, is an indicator of whether properties are seen as essential to the category, and are therefore expected to be found in all instances, then all generic subjects except the DS seem to be on par with merely statistical knowledge. Looking at the specific data points for the DS, there is a tendency towards a bimodal distribution, with 11 participants' scores between 1-5 points, and the remaining three between 9-12 points.

Based on the literature, this indicates that a DS subject allows for kind reference, but that it is more ambiguous between the individual and the kind reading than the other subject types in this context. Future studies with a larger sample size will be helpful in further addressing this hypothesis. These should further consider whether the types of questions asked in the test booklet are appropriate to investigate whether a connection between kind and property is likely to be principled. A re-evaluation thereof might explain the similar responses for category-property scores for BP and IS and essentialism scores for all three generic subjects in this experiment.

These results call for two comments: First, looking at the results in Section 5.2.4, a more fine-grained analysis will aid our understanding and interpretation of these results, especially with regard to the DS subjects. Second, previous literature provides strong support for the hypothesis that generic language strengthens the link between category and property more than mere prevalence (and vice versa leads to higher prevalence estimates than typically found, e.g. Khemlani et al., 2007, Leslie et al., 2011). Thus, even though MQ subjects seem to fare as well as generic subjects, there is a difference in how they are processed.[15]

[15] Returning to some of the issues raised at the start of this chapter, running a version of this study that collects and analyses reaction times would be beneficial in addressing this question and its relation to *Generics-as-Default* in more detail.

Essentialism The tasks in the essentialism composite targeted properties that were expected to be highly prevalent amongst members of the kind. If they are essential to the category, it should be possible to extend them to other instances of the kind. The specific questions in the test booklet that targeted the essentialist beliefs of the participants addressed 'explanation (content)', 'stability', 'inheritance', 'novel induction', and 'match (taxonomic vs shape vs property)' as outlined in Table 5.1 above. When answering these questions, properties that are accidental to the kind would not be sufficient to allow for the extension of these properties to other instances. To illustrate, in the 'inheritance' task, a novel *zarpie*, as a member of the kind, would still possess properties that other *zarpies* possess, even if it did not grow up amongst other instances of the kind. It is expected to have its inherent zarpie-like properties unless there is a process or mechanism that explains its lack thereof (similar to how albino tigers have a different colour pattern due to an absence in pigmentation). Similarly, a property tested for 'stability' would be expected to be present in one individual zarpie, regardless of its age.

A oneway ANOVA was conducted to establish whether the differences in responses (by participant) for the questions of the composite in the test booklet specifically targeting expectations of essentialism based on the 4 wording conditions (wording: BP vs IS vs DS vs 'most', between participants) were significant. This analysis indicated no significant effect of wording condition on the scores in the essentialism composite ($F(3,52) = 0.3996$, $p = .7539$), as plotted in Figure 5.5.

Due to lack of interactions in the omnibus model, no paired comparisons were carried out.

The four subjects all score very similarly in the totals of this component: the BP scores 63.10%, the IS scores 63.69%, the DS scores 60.71%, and the MQ scores slightly lower at 58.33% (see Figure 5.5). For comparative purposes, the essentialism scores from Gelman et al. (2010) for their control condition 'this' are included in this graph.[16]

[16]No error bars are included for that condition as the full data set was not available, but the full analysis of their data is available in the original article.

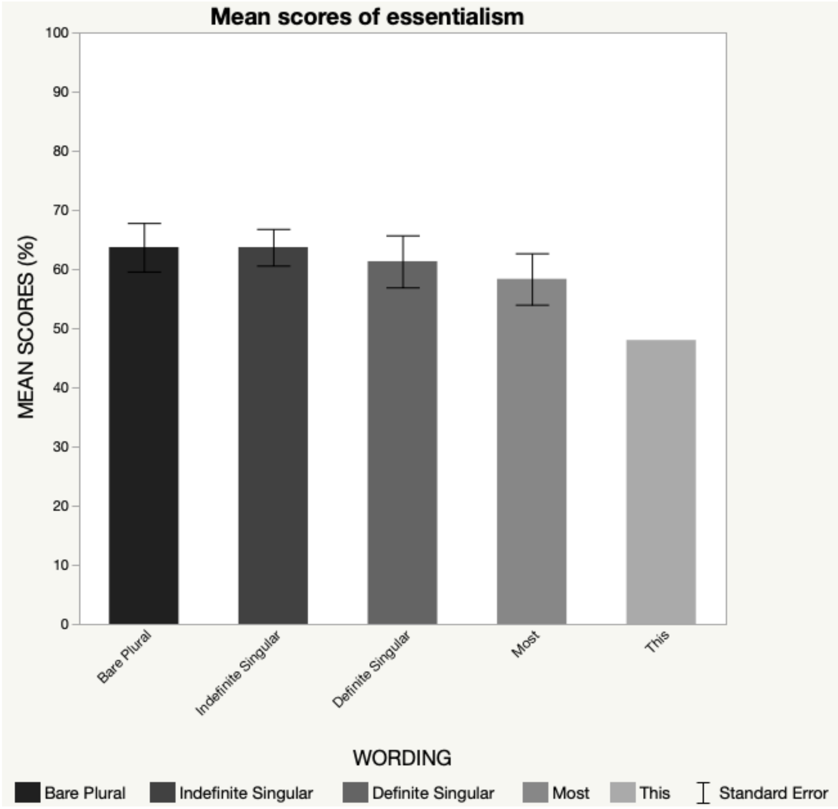


Figure 5.5: Mean scores of Essentialism Composite (Experiment 13 (narrative)), by wording. Error bars = ± 1SE

These results show that the prediction that the IS implicates the properties are more essentially connected do not come out in these results. This is confirmed by independent findings in Rhodes et al. (2012), following up on Gelman et al. (2010), where no differences between BP and IS generic subjects were found. The reason this experiment included the third generic subjects was to, first and foremost, replicate the previous findings as well as compare them to the DS subject, which has not been explored. The exclusion of these other experimental conditions in these prior studies might be due to its lack of prominence in everyday speech. Further, based on previous pseudoword studies, there is evidence that the morphosyntax of the subject matters with regard to the expectations raised. One explanation why the results in this study were less pronounced is that only animal categories were tested. These are known to be stable and more essentialist than social kinds or artefacts and as such, the experiment might be subject to ceiling effects, which do not allow for subtle differences between the subject types to come out.[17]

Memory Participants' memory recall rate was lowest for DS subjects (67.86%), while BP, IS, and MQ conditions had a more accurate recall, at 76.79%, 73.21%, and 78.57%, respectively (see Figure 5.6, note that again the data for the specific control condition 'this' from Gelman et al., 2010 were included for comparative purposes). For DS subjects, many participants falsely recalled the subject as a demonstrative, which only allows specific reference (e.g. 'this zarpie'). This supports the explanations of the lower category-property links scores by providing further evidence that certain participants did not have a generic interpretation of the information presented.

A oneway ANOVA was conducted to establish whether the differences in responses (by participant) for the questions of the composite in the test booklet specifically targeting memory recall based on the 4 wording conditions (wording: BP vs IS vs DS vs 'most', between participants) were significant. This analysis indicated no significant effect of wording condition on the scores in the memory composite ($F(3,52) = 0.56$, $p = .6438$), as plotted in Figure 5.6.

Due to lack of interactions in the omnibus model, no paired comparisons were carried out.

Notably, participants scoring high on the category-property link composite and the essentialism component of the study recalled the subject form more accurately or sometimes even falsely recalled them as another generic subject type (e.g. as *zarpies*). One might argue that this is a recall effect based on the introductory prompt for each trial, regardless of the subject condition of the picture book, since every prompt began by stating "Look at this zarpie! The/A/This zarpie(s) ...".[18] This hypothesis has not specifically been tested experimentally. It is possible that this prompt has a

[17] The essentialism questionnaire, assessing the participants' baseline essentialist beliefs, was marginally, but not significantly, higher for the IS group than the others. It might be worthwhile to investigate the post-experiment effect on essentialist beliefs. Just like in priming studies, it is possible that participants' general beliefs about essentialism were affected by the the effect of the language used during the experiment.

[18] I would like to thank anonymous reviewer, who commented on a draft for a paper presenting these studies in Fuellenbach & Gelman (2019), for pointing out these alternative explanations.

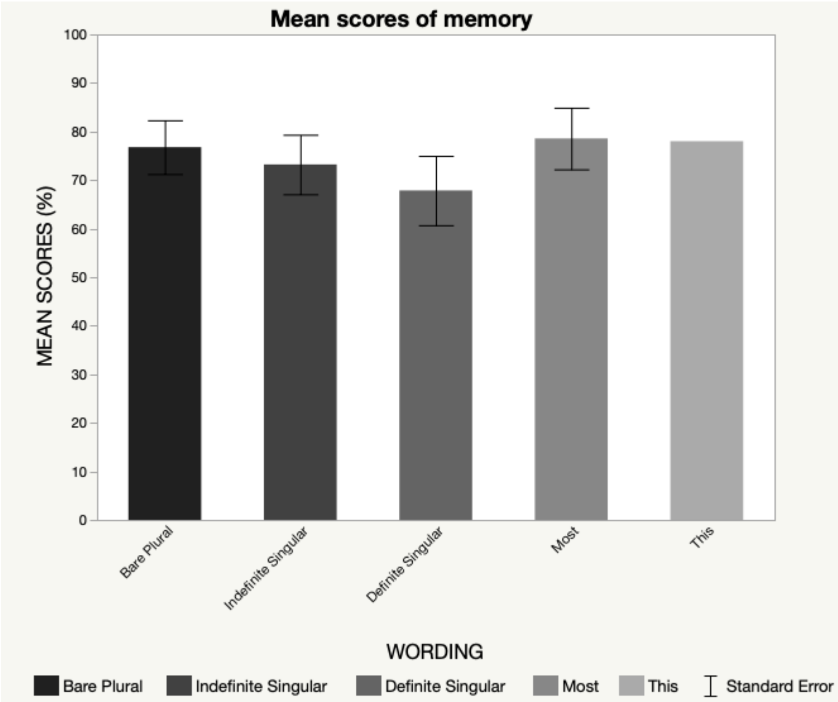


Figure 5.6: Mean scores of Memory Composite (Experiment 13 (narrative)), by wording ('BP' = bare plural, 'IS' = indefinite singular, 'DS' = definite singular, 'MQ' = most quantified).
Error bars = ± 1 SE from the mean.

carry-over effect on the recall of the generic (or control) sentence. However, this did not affect the memory recall for any of the other wording conditions, although one might argue that the demonstrative 'this' and the definite determiner 'the' have the largest phonological overlap and this effect could be hypothesised to be largest in this condition. Further, this brings forward a discussion on whether memory recall is a useful proxy to ensuring whether distinct generic interpretations were achieved based on the morphosyntactically distinct subject forms. While I believe that memory components are a useful way to check attention, I do not think that the implicit learning that was targeted in this study is dependent on explicit recall of the linguistic form, and it was not explicitly assumed to be a prerequisite in the pseudoword studies, nor did I find any correlation in the *kevtas* studies.

It could be of interest to investigate these differences with more detailed memory tests in future studies to better understand when DS subjects are recalled as generic and when they are misremembered as specific. In that case, the memory test should be better integrated in the experimental design and constructed in a way that goes beyond an attention check to ensure that the results would allow us to make more informed assumptions about the potential patterns in misremembered generic subjects.

Similarly, some of the responses in the IS condition falsely recalled the subject form as a plural, indicating that they had a generic interpretation but did not specifically recall its linguistic form. One the one hand, this suggests that the subtleties of morphosyntactic form can influence our conceptual interpretation without the need for correct recall of its form. These falsely remembered subject forms hint at how participant encode conceptual input but do not rely on the exact form in their own output, if another form with a similar function can also be used. On the other hand, these false recalls can also be seen as problematic, in that they might be stored without the associated semantic representations that distinguish them from one another.

5.2.3.6 Discussion of composite results

In general, this replication of Gelman et al. (2010) confirmed the role of generic language in establishing tight links between categories and properties and that the paradigm of the original *zarpies*-study may partially extend to other generic subjects as well. However, the test batteries of the original studies were not able to confirm the distinct hypotheses for the three separate generic subjects in English. While there was a difference between generic and quantified language, there were no interactions between the three generic subjects. Even though some subject types were rated marginally higher for category-property links, overall, these differences were small, and not found for the essentialism composite.

Moving forward, two choices were made regarding further studies and analyses. Firstly, if these distinctions cannot be targeted in competent adult speakers, the likelihood of finding differential results in child participants is very low. Thus, this specific study was not extended to include child participants. Secondly, during the main analysis, I found that some of the items were trending towards the hypothesised predictions more so than others. These higher scores for category-property links or essentialism scores occurred across the board, independent of the wording condition they were in. Therefore, I have decided to conduct an item-analysis to explore the effects of specific questions of the test booklet, which is presented in Section 5.2.4 below.

5.2.4 Post-hoc item analysis

Experiment 13 (narrative) showed that the paradigm of the original *zarpies* study may extend to singular generic subjects. However, the questions of the test booklet did not show distinct behavioural responses that supported the hypotheses for the three generic subjects in English. Focusing on the effects of specific items rather than category-property links by wording conditions illustrates that some properties were consistently rated higher than others, indicating stronger category-property links across the board. Therefore, I conducted an item analysis to disambiguate the effects of wording from the effects of certain properties predicated of animal kinds in particular.

I looked at the effects of specific items used in the test booklet which might allow for a better analysis and linking back to the findings of the initial research hypothesis. This specifically concerns the conceptual connection of properties and their connection type to a category, proposed by Prasada & Dillingham (2006, 2009), which goes beyond testing the strength of a category-property link for a novel animal. This also returns to the distinction of property types in footnote 11 above.

For this analysis, I introduced an alternative scoring method in addition to the one in Gelman et al. (2010). The score of each property within a battery of questions, i.e. within the category-property composite and within the essentialism composite, was compared to the scores this property received in the other wording conditions. This analysis confirmed the hypothesis that certain properties scored reliably higher than others across the board, regardless of the wording of the subject. Therefore, the high scores are unlikely to be attributable to the linguistic form but rather reflect something characteristic of the property itself. It seems that something about the property itself appears to make it easily extendable to new instances, accounting for higher scores of questions targeting category-property links and essentialism.

The following sections show an analysis that uncovers high variation of acceptance rates within the category-property composite and the essentialism composite. Looking at the means across all properties and conflating over the different wording conditions allowed to demonstrate patterns for properties that are obscured by not looking at the property's scores individually, but grouped by subject form. The individual properties that were assessed in the testing phase of the experiment were compared first collapsed by wording condition, and then afterwards separated by wording. Note that the latter, more fine-grained separation, results in very few data points per condition and should therefore be taken as an exploratory analysis that serves as a pilot study for future experiments.

5.2.4.1 Familiar property induction

I begin by looking at the familiar property induction task (as introduced in Table 5.1 above), which targets category-property links by asking whether participants can "find another one" based on a property they have encountered before. The totals by wording are displayed in Figure 5.7.

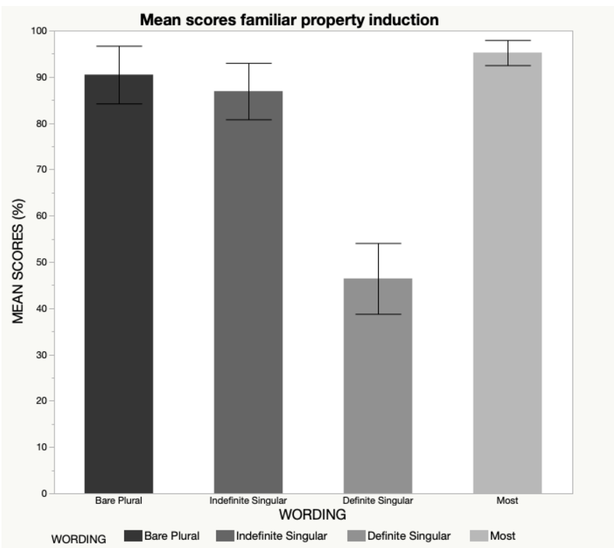


Figure 5.7: Item analysis for familiar property induction task. Error bars = ± 1SE.

The results from just the properties that were used in the familiar property induction task illustrate the systematically lower scores for DS subjects. Table 5.3 provides an overview of all the scores for the familiar property induction task.

Table 5.3: Summary of item analysis (familiar property induction)

WORDING CONDITION	**TOTAL(%)**
Bare Plural	90.48
Indefinite Singular	86.90
Definite Singular	46.43
most-quantified	95.24

Familiar properties by item Collapsed over all four wording conditions, i.e. just looking at the patterns of the properties, we find that ‘stands up on its hind legs when surprised’ scores highest and ‘bounces a ball on its back’ scores lowest.

Post-hoc paired comparisons investigated the differences between the three properties ‘bounces a ball on its back’, ‘stands up on its hind legs when surprised’, and ‘hates ice-cream’. A oneway ANOVA was conducted to establish whether the differences in responses for these three properties were significant. This analysis detected no significant difference of the scores for the different properties ($F(2,165) = 2.3214$, $p = .1013$). This is illustrated in Figure 5.8.

Recall that the properties that were used in the familiar property induction task show that the only systematic difference by wording is for DS subjects, which score much lower than the other wording conditions. The totals for each wording condition were summarised in Table 5.3 in Section 5.2.3.5 on page 240.

While there are no differences between the three properties when collapsed by wording, I will now turn to an analysis of the properties, separated by wording condition.

Familiar properties by item by wording The following is a preliminary and exploratory discussion of the individual data patterns of the properties within each wording condition. It indicates that there could be distinct relative patterns

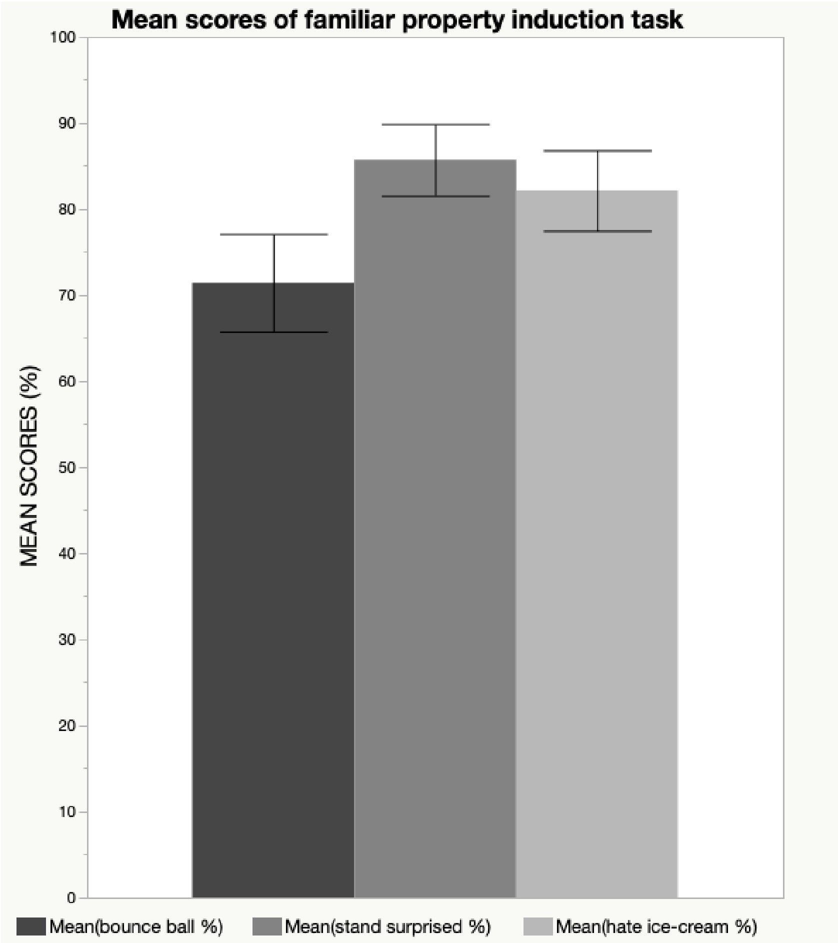


Figure 5.8: Item analysis for familiar property induction task by item. Error bars = ± 1SE.

for each property, emerging within each wording condition. The differences in this analysis are all at different magnitudes. Due to the small sample size, this discussion of the effect of wording on each individual property type introduced in the booklet does not have enough power to generalise based on these results. This and the following explorations looking at the composites by item by wording are all to be treated as exploratory analyses.

Similar to the by-item-analysis in the previous paragraph, Figure 5.9 shows that relatively speaking, 'bounces a ball on its back' scores lower in general than the other properties, and that 'stand up on its hind legs when surprised' scores higher or equally high than the other properties. I will provide a more general discussion of the possible underlying reasons for this specific distribution amongst the properties in Section 5.3 below.[19]

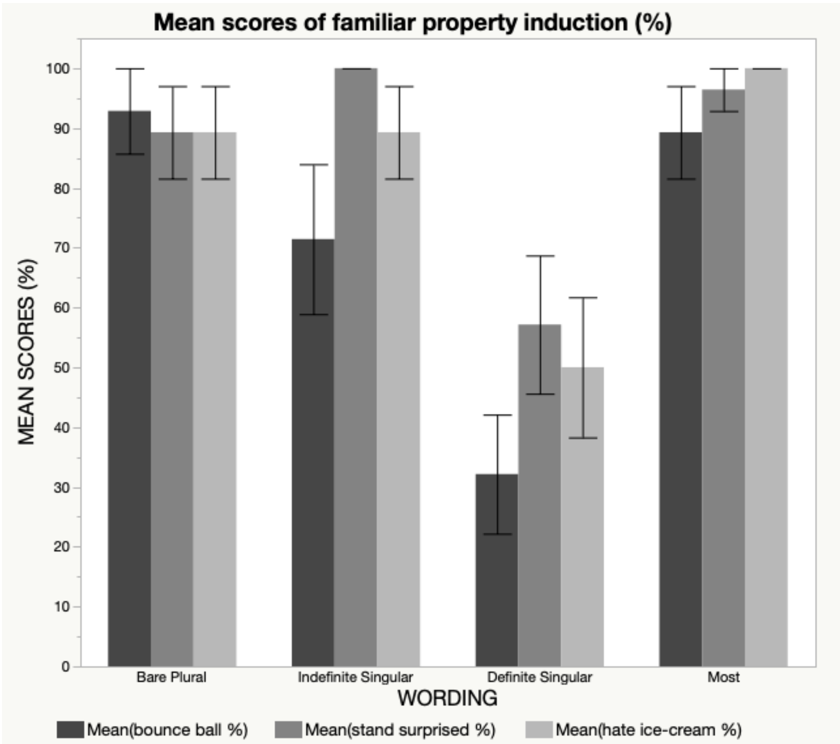


Figure 5.9: Mean scores for familiar property induction task by item by wording. Error bars = ± 1SE.

The pattern for the by-item analysis mirrors the findings in the general analysis laid out in paragraphs **Category-property links** through **Memory** in Section 5.2.3 above. Moreover, the variation present within the IS subjects shows that the property 'bounces a ball on its back' (see Figure 5.7) is lower than the other conditions, the same pattern as for the DS subjects, even if at an overall smaller magnitude[20] (see Table 5.4).

By contrast, the property 'stands up on its hind legs when surprised' scores highest for the IS subject. For the BP and *most*-quantified condition, there is not as much variation between properties (see Table 5.5). The DS has the highest SD and sample variance.

[19] All of these item-analyses show a significant difference between the ratings of the properties, grouped by subject type, which is always driven by low scores for DS subjects. No differences emerge between the other subject types. However, due to the low number of observations per property by wording conditions, no full statistics are reported here as they are currently underpowered. Instead, descriptive statistics are provided for each of the properties, separated by wording.

[20] In the *most*-quantified condition, the property 'bounce a ball on its back' is also the lowest

Table 5.4: Summary of familiar property induction ("bounces a ball on its back")

bounce_ball	**BP**	**IS**	**DS**	**MQ**
Mean	1.86	1.43	0.64	1.79
Standard Error	0.14	0.25	0.20	0.15
Median	2.00	2.00	0.50	2.00
Standard Deviation	0.53	0.94	0.74	0.58
Sample Variance	0.29	0.88	0.55	0.34

Table 5.5: Summary of familiar property induction ("stands up on its hind legs when surprised")

stand_surprised	**BP**	**IS**	**DS**	**MQ**
Mean	1.79	2.00	1.14	1.93
Standard Error	0.15	0.00	0.23	0.07
Median	2.00	2.00	1.00	2.00
Standard Deviation	0.58	0.00	0.86	0.27
Sample Variance	0.34	0.00	0.75	0.07

The third property tested within the familiar induction paradigm was 'hates ice-cream'. It scores equally high for BP and IS subjects, at 89.29%, and is on par with the other properties within BP subjects and between the other two within IS subjects. In terms of relative scoring to the other properties, the same pattern as for the IS subjects emerges within DS subjects (see Table 5.6).

one but the difference between this property and the other three is not as pronounced.

Table 5.6: Summary of familiar property induction ("hates ice-cream")

hate_ice-cream	**BP**	**IS**	**DS**	**MQ**
Mean	1.79	1.79	1.00	2.00
Standard Error	0.15	0.15	0.23	0.00
Median	2.00	2.00	1.00	2.00
Standard Deviation	0.58	0.58	0.88	0.00
Sample Variance	0.34	0.34	0.77	0.00

5.2.4.2 Novel property induction

The same kind of analysis, with the same caveats, was conducted for the properties that targeted novel property induction. In the test booklet, participants were asked about three properties that they had not encountered in the picture book. These were: 'dances in circles', 'buzzes when angry', and 'plays frisbee' (Figure 5.10). They then judged in terms of whether they think another zarpie, again one they had not seen in the picture book, would also have this property.

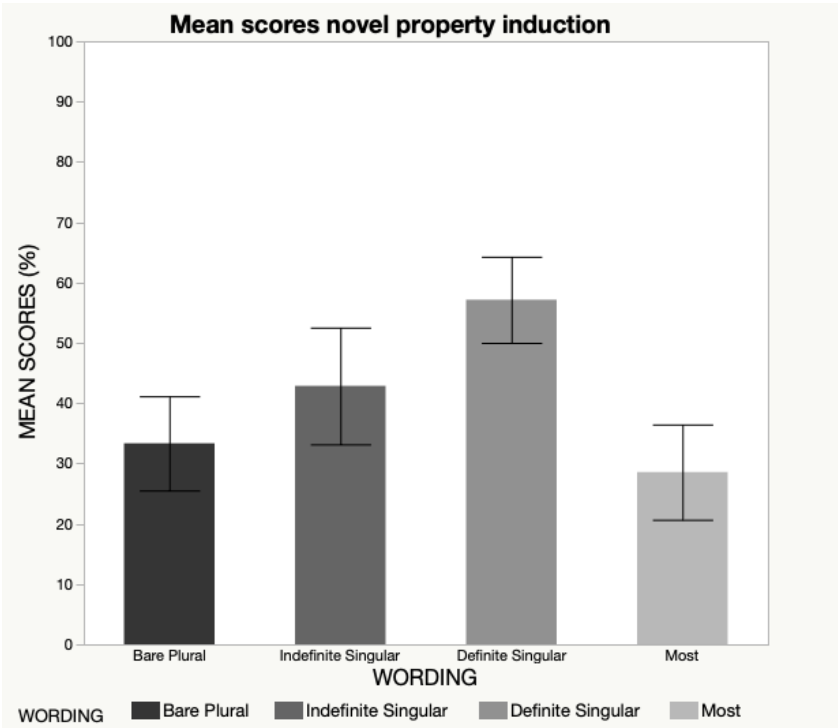


Figure 5.10: Mean scores for novel property induction task. Error bars = ± 1SE.

An overview of the results from the properties that were used in the novel property induction task is provided in Table 5.7, which summarises the total scores.

Table 5.7: Summary of item analysis (novel property induction)

WORDING CONDITION	TOTAL(%)
Bare Plural	33.33
Indefinite Singular	42.86
Definite Singular	57.14
most-quantified	28.57

The pattern of the totals of the novel property induction is different from the familiar property induction task. Most significantly, both BP and MQ score low while IS and DS score higher. While the highest score for the familiar property induction task was found in the *most*-quantified condition and the lowest score in the DS condition, this pattern is reversed for the novel property induction task. Thus, when recalling a previously introduced property, the DS subject does not raise the expectation that another instance has said property. If for a previously introduced property participants were told that most instances have it, they are more likely to expect that a novel instance in fact possesses that property. However, with a property that was not introduced before, as in this novel property induction task, participants do not expect that most instances possess the new property.

Novel properties by item Collapsed over the four wording conditions, i.e. just looking at the patterns of the properties, 'dances in circles' scores lower than both 'buzzes when angry' and 'plays frisbee', which score similarly.

Post-hoc paired comparisons investigated the differences between the three properties 'dances in circles', 'buzzes when angry', and 'plays frisbee'. A oneway ANOVA was conducted to establish whether the differences in responses for these three properties were significant. This analysis indicated a significant difference for the scores for the different properties ($F(2,165) = 6.0498$, $p = .0029$). This is illustrated in Figure 5.11.

Paired comparisons show that, compared to 'dances in circles', the property 'buzzes when angry' differs significantly ($p = .0071$), and so does 'plays frisbee' ($p = .0099$). There are no differences between 'buzzes when angry' and 'play frisbee' ($p = .9934$).

Overall, all three properties received lower scores than the familiar properties in the previous task. I hypothesise that this might be due to the unfamiliarity and therefore provides support for the idea that the training phase with the picture book helped participants understand the tight link between the property that is predicated of the zarpies.

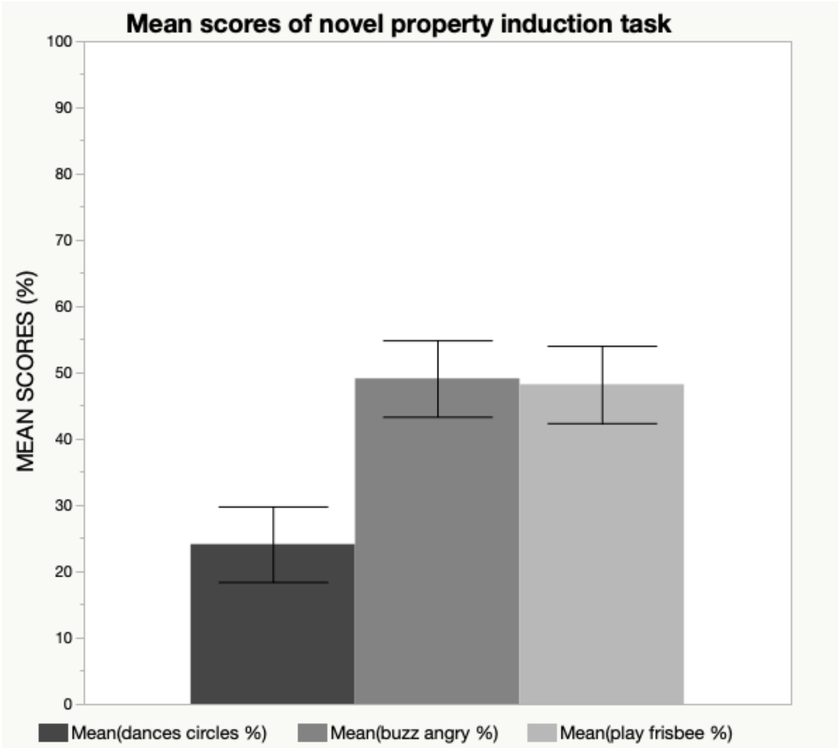


Figure 5.11: Mean scores for novel property induction task by property. Error bars = ± 1SE.

Table 5.8: Summary of novel property induction ("dances in circles")

dance_circles	**BP**	**IS**	**DS**	**MQ**
Mean	0.79	0.43	0.43	0.29
Standard Error	0.26	0.23	0.23	0.19
Median	0.00	0.00	0.00	0.00
Standard Deviation	0.97	0.85	0.85	0.73
Sample Variance	0.95	0.73	0.73	0.53

Novel properties by item by wording Parallel to the familiar property item analysis, I had a preliminary look into the patterns for the scores of different properties within each of the four wording conditions. Again, the current study is underpowered to conduct a full analysis, but the same tendency towards repeated relative patterns for IS and DS subjects became apparent.

Focusing again at the patterns of each property, broken down by wording condition, novel properties patterned as shown in Figure 5.12.

Compared to the other novel properties, 'dances in circles' (descriptive statistics

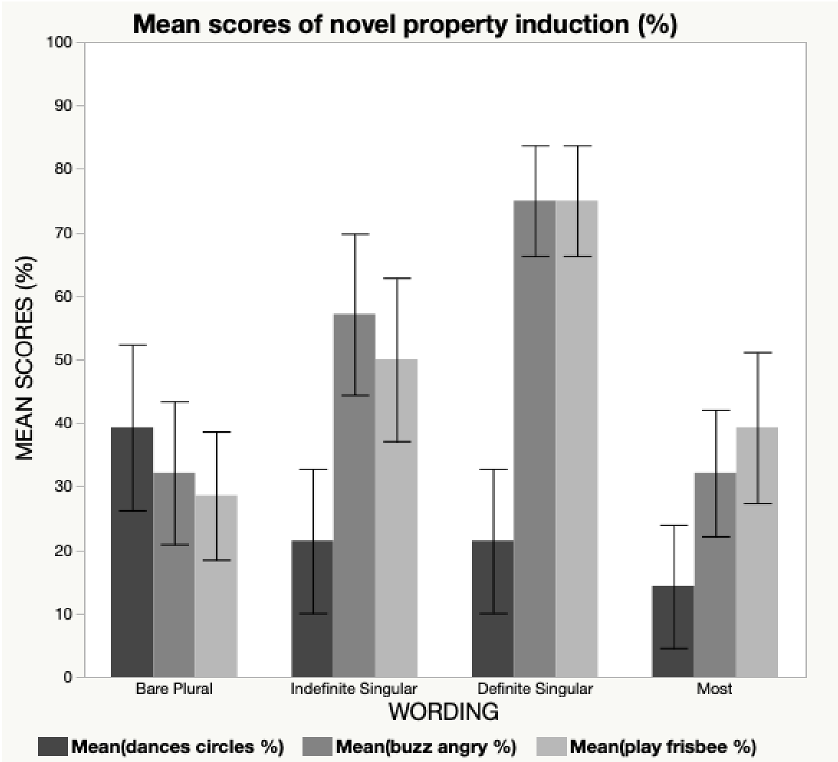


Figure 5.12: Item analysis for novel property induction task by item by wording. Error bars = ± 1SE.

in Table 5.8) scores the highest when presented with a BP subject. One account for this pattern is that this property can easily be conceptualised as an idiosyncratic property, as the specific style of dancing might easily vary between members of a kind, analogous to existing categories in the world. This could be due to the fact that properties that are introduced with a generic BP subject, even if only statistically related, are expected to be present in most instances of a kind.

By contrast, 'buzzes when angry' scores high for IS and DS subjects. The physiological response of buzzing appears less idiosyncratic and more easily generalisable across instances. This type of characteristic can be interpreted as more likely to be principally connected to the kind, analogous to dogs barking, and therefore scores higher with IS and DS subjects, which are linguistically associated with properties that are, compared to the BP, more likely to be principally connected and less likely to be statically connected, respectively. For summary statistics, see Table 5.9.

The third novel property that was tested was 'playing frisbee' (see Table 5.10).

Table 5.9: Summary of novel property induction ("buzzes when angry")

buzzes_angry	**BP**	**IS**	**DS**	**MQ**
Mean	0.64	1.14	1.50	0.64
Standard Error	0.23	0.25	0.17	0.20
Median	0.00	1.50	2.00	0.50
Standard Deviation	0.84	0.95	0.65	0.74
Sample Variance	0.71	0.90	0.42	0.55

Table 5.10: Summary of novel property induction ("plays frisbee")

play_frisbee	**BP**	**IS**	**DS**	**MQ**
Mean	0.57	1.00	1.50	0.79
Standard Error	0.20	0.26	0.17	0.24
Median	0.00	1.00	2.00	0.50
Standard Deviation	0.76	0.96	0.65	0.89
Sample Variance	0.57	0.92	0.42	0.80

Interestingly, the pattern resembles the one of 'buzzes when angry' in that the IS and DS scored highest, with the DS scoring equally high for both properties (although note that I do not see any obvious reason for the resemblance of these two properties). The BP, however, scored lower than for both other novel properties.

5.2.4.3 Inheritance task

I now turn to the questions targeting beliefs about inheritance of properties. Recall that participants were asked about four properties and whether another zarpie would display those, inherited as a member of the kind zarpie, or whether they would have an alternative property, acquired as a result of being raised by another kind of animal. The totals by wording are displayed in Figure 5.13.

The results are very similar across all wording conditions, as illustrated in Table 5.11.

Inheritance task by item Post-hoc paired comparisons, collapsing over the four wording conditions, investigated the differences between the four properties

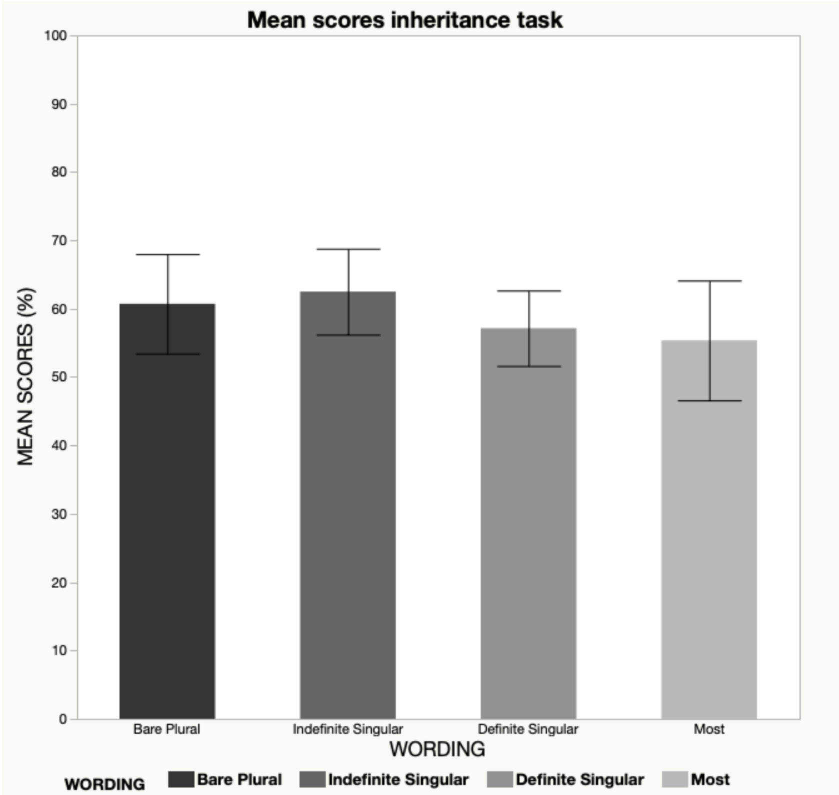


Figure 5.13: Mean scores for inheritance task. Error bars = ± 1SE.

Table 5.11: Summary of item analysis (inheritance task)

WORDING CONDITION	TOTAL(%)
Bare Plural	60.71
Indefinite Singular	60.71
Definite Singular	57.14
most-quantified	55.36

in the test booklet: 'chases shadows', 'sings', 'wiggles its horns when it is happy', and 'eats tulips'.

A oneway ANOVA was conducted to establish whether the differences in responses for these four properties were significant. This analysis indicated a significant difference for the scores for the different properties ($F(3,220) = 18.4391$, $p < .0001$). This is illustrated in Figure 5.14.

Paired comparisons show that, compared to 'chases shadows', all other properties differ significantly (vs 'wiggles its horns when it is happy' $p = .0001$, vs 'eats tulips' p = .0001, vs 'sings' p = .0009). Further, 'wiggles its horns when it is happy' and 'sings' also differ significantly (p = .0086). There are no differences between the other properties ('wiggles its horns when it is happy' vs 'eats tulips' p = .4436; 'sings' vs 'eats tulips' p = .323).

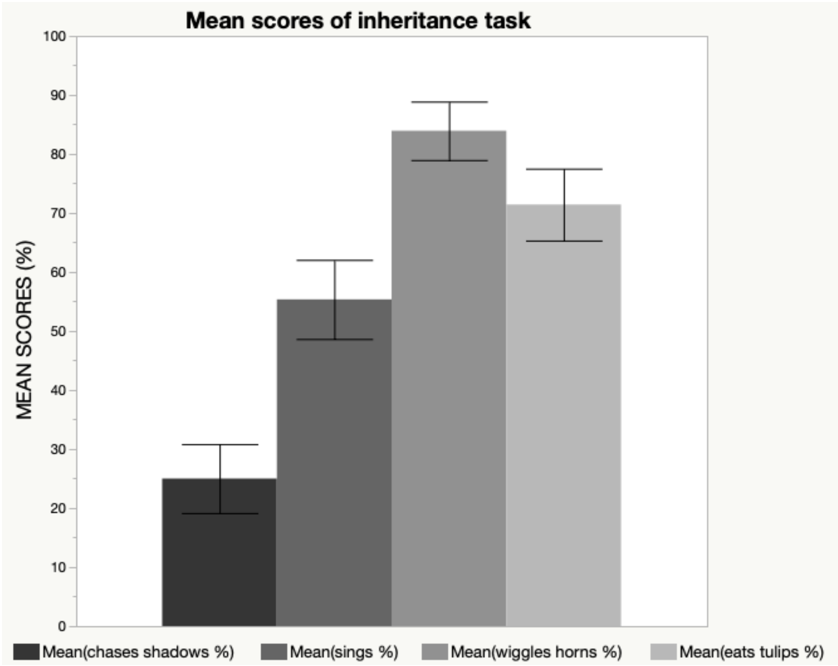


Figure 5.14: Mean scores for inheritance task by property. Error bars = ± 1SE.

Inheritance task by item by wording Once again, I had a preliminary look into the patterns for the scores of the four different properties within each of the four wording conditions. Again, the current study is too underpowered to conduct a full analysis and I will only provide descriptive statistics for each of properties.

Looking at the patterns of each property, broken down by wording condition, the patterns illustrated in Figure 5.15 emerge.

As for the familiar property induction tasks and the novel property induction tasks, the distributions for each of the properties are presented in separate tables below. Table 5.12 shows that for 'chases shadows', if introduced with a DS subject, no participants predicted that a zarpie raised by non-zarpies would also chase shadows. By contrast, if introduced with an IS, almost half of the participants extended the property in the inheritance question. BP and MQ subjects fared

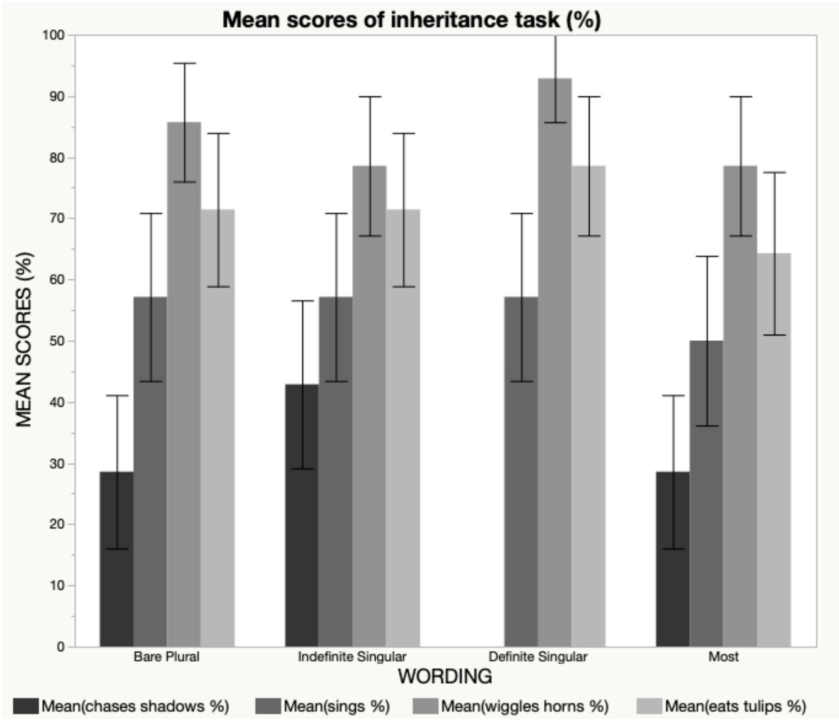


Figure 5.15: Item analysis for inheritance task by property by wording. Error bars = ± 1SE.

equally with just under one third of the participants predicting that 'chases shadows' would be an inherited trait.

Table 5.12: Summary of inheritance task ("chases shadows")

chases shadows	**BP**	**IS**	**DS**	**MQ**
Mean	0.29	0.43	0.00	0.29
Standard Error	0.13	0.14	0.00	0.13
Median	0.00	0.00	0.00	0.00
Standard Deviation	0.47	0.51	0.00	0.47
Sample Variance	0.22	0.26	0.00	0.22

The second property which participants had to judge as being inherited was whether a zarpie likes to 'sing'. Table 5.13 shows the descriptive statistics for this property. All subject types were rated very similarly, as visualised in Figure 5.13.

The third property, 'wiggles its horns when it is happy', rendered the highest scores of all the properties tested for inheritance. All means were equal to or higher than 79%, with the DS receiving the highest scores (Table 5.14).

Table 5.13: Summary of inheritance task ("sings")

sings	**BP**	**IS**	**DS**	**MQ**
Mean	0.57	0.57	0.57	0.50
Standard Error	0.14	0.14	0.14	0.14
Median	1.00	1.00	1.00	0.50
Standard Deviation	0.51	0.51	0.51	0.52
Sample Variance	0.26	0.26	0.26	0.27

Table 5.14: Summary of inheritance task ("wiggles horns")

wiggles horns	**BP**	**IS**	**DS**	**MQ**
Mean	0.86	0.79	0.93	0.79
Standard Error	0.10	0.11	0.07	0.11
Median	1.00	1.00	1.00	1.00
Standard Deviation	0.36	0.43	0.27	0.43
Sample Variance	0.13	0.18	0.07	0.18

Finally, the fourth property tested was 'eats tulips'. Table 5.15 shows that this property also received high scores across the board. Again, the DS received the highest mean score out of the four subjects, and the MQ subjects had the lowest average score.

Table 5.15: Summary of inheritance task ("eats tulips")

eats tulips	**BP**	**IS**	**DS**	**MQ**
Mean	0.71	0.71	0.79	0.64
Standard Error	0.13	0.13	0.11	0.13
Median	1.00	1.00	1.00	1.00
Standard Deviation	0.47	0.47	0.43	0.50
Sample Variance	0.22	0.22	0.18	0.25

Summary of analysis by property type, collapsed over wording condition

In general, these data show that there are some patterns that show up for the different properties. More importantly, this trend holds regardless of the wording condition in which they were presented, as this analysis collapsed over wording condition. The amount of data collected in this pilot study does not give sufficient power to analyse whether these patterns are significant. However, looking at the preliminary data patterns displayed by specific properties that consistently scored higher, independent of their subject form, raises the question whether there is an alternative explanation.[21] It is noteworthy that these specific items were properties of the type 'wiggles its horns', 'eats tulips', 'buzzes when angry', and 'stands up on hind legs'. These properties can be argued to be properties that are typically thought to be part of *overhypotheses* in animal categories. In short, these are the types of properties related to locomotion, diet, reproduction, etc., which are more easily generalised in animal categories as they are thought to be more consistent across different kinds and subkinds. The next section relates this theory to the patterns observed in this section in more detail.

5.3 Overhypotheses

This section presents an alternative theory of explaining how certain properties are expected to be connected to novel kinds. The aim is to provide a more comprehensive analysis of the responses for the category-property links found in Experiment 13 (narrative) in Section 5.2.3. As the by-item analyses illustrated, there was wide variation in how some specific properties were perceived to be linked to the *zarpies*, although the variation itself was fairly consistent between participants. By considering this alternative theory, I will not only provide an alternative explanation for the specific observed variation in this study, but also integrate the findings with higher level theories. Consequently, this allows predictions about the behaviour of specific types of properties, predicated of animal kinds, for future experiments. This could be either follow-up studies designed specifically to hone in on the relationship between the various properties attributed to *zarpies*, or studies that more generally research the interaction of generic language and properties predicated of animal kinds.

[21] No differences for individual items were detected in the taxonomic vs thematic vs shape task. For all three generic subjects, the taxonomic category was the most chosen one. This contrasts with the MQ subjects, where shape was equally likely to be chosen (39.39% for taxonomic and 35.71% for shape choice, with 21.43% for thematic choice). Since this task is not strongly related to the research question at hand, it was disregarded in this item analysis.

The most promising view to account for the observed variation in the category-property links is the theory of overhypotheses. The notion of overhypothesis was first introduced by Goodman (1955) (cited from Goodman, 1983, Chapter IV), addressing a variation of the problem of induction (Hume, 1748). Goodman argues that there are multiple levels involved when abstracting away from specific instances to arrive at a generalisation. In order to capture similarities from a variety of specific instances, we generate these abstract hypotheses. In so doing, we focus less on the differences between kinds and subkinds, and more on shared basic properties across members of the same superordinate kind.[22]

The reason why forming overhypotheses is so appealing from a cognitive viewpoint is that they offer a quick and efficient form of inductive learning. Specifically, they provide "a second-order generalization about categories in general" (Dewar & Xu, 2010, p. 1871). It is this form of generalisation that allows us to form hypotheses not only about one category that we are exposed to, but all categories that are similar to each other and related via a superordinate kind. Goodman (1983) provides an example from outside the realm of natural kinds (which Experiment 13 (narrative) has focused on). Instead, in Goodman (1983)'s thought experiments, participants were shown a marble from a bag containing multiple marbles. Each marble from a new bag varies in colour but crucially, every bag only contained marbles of one colour. A first-order hypothesis about a novel bag would be to infer the colour of the other marbles to be the same as the one the participant is shown. Forming a second-order generalisation, i.e. an overhypothesis, means the participant assumes that bags of marbles are always uniform in colour. Kemp et al. (2007, p. 308), provide a broader and more recent overview of second-order generalisations (and contextualise it within a Bayesian model). Regarding the design or Goodman (1983)'s original experiments, they elaborate:

[22] Specifically how members are understood to belong to the same superordinate kind is not addressed in Goodman, especially not members of animal kinds. However, given the cognitive biases that I have explored throughout this thesis, it seems safe to assume that *shape* and *taxonomic* biases play a significant role. This is because we can expect the same mechanism to apply when deciding which basic kinds are members of a superordinate kind that we expect to inform categorisation of individuals to be members of a kind.

> Suppose that S is a stack containing many bags of marbles. We empty several bags and discover that some bags contain black marbles, others contain white marbles, but that the marbles in each bag are uniform in color [sic]. We now choose a new bag - bag n - and draw a single black marble from the bag. On its own, a single draw would provide little information about the contents of the new bag, but experience with previous bags may lead us to endorse the following hypothesis:
>
> **H**: All marbles in bag n are black. If asked to justify the hypothesis, we might invoke the following overhypothesis:
>
> **O**: Each bag in stack S contains marbles that are uniform in color.

The alternative hypothesis is that upon seeing bags containing marbles of different colours, that each bag contains marbles of these possible colours. As Dewar & Xu (2010, p. 1871) explain, "[t]he first-order generalization concerns the contents of each individual bag (...). The second-order generalization, or overhypothesis, is that 'bagfuls of marbles are uniform in color'."

Most learning is dependent on generalisations from individual exposures and exposure to individuals. However, this cannot possibly account for the variety and variation of properties in the real world (e.g. Quine, 1960). Within the context of second-order generalisations, this becomes more obvious seeing that not all learning leads to the formation of an overhypothesis. In fact, it is not yet clear which categories allow for overhypotheses to be formed. To address this question, some experimental work has started to investigate this phenomenon in more detail and has repeatedly shown that at least certain properties of animal kinds are susceptible to overgeneralisation. Particularly relevant for the topic of this thesis, overhypotheses have been argued to be useful to account for the quick acquisition of conceptual knowledge in the developmental literature (e.g. Marchak & Gelman, 2018). Marchak & Gelman argue that acquiring a new concept and forming hypotheses about its properties typically occurs after only very little exposure to very few instances. In these instances, the function of overhypotheses would be to fill in the gaps in specific knowledge about a particular kind by providing general information based on the occurrence of properties in this kind based on what is known about the superordinate kind. This is in line with the findings of the studies in this chapter and the previous ones in this thesis.

Taking into account this independently developed idea of second-order generalisations, it looks likely that the notion of overhypothesis can also be used to explain patterns in the experimental data for the *zarpies* study. The main reason for this lies in the observation that the generalisations made about the animal kind varied systematically based on the type of property predicated of them, and specifically not on the type of subject form with which the *zarpies* were introduced to participants. However, just noting that some properties systematically scored higher than others is not sufficient to verify whether overhypothesis offers a valid interpretation for these results. I will therefore now jointly look at the analysis of the post-hoc item analysis from Section 5.2.4 and the main characteristics of overhypothesis.

5.3.1 Overhypothesis for *zarpies*

Recall that for the post-hoc item analysis, an alternative scoring method was used, which allowed an assessment of how individual properties fared across wording conditions. The types of properties that scored high for category-property links across the board, regardless of the wording of the subject, are those that are easily hypothesised to be principally connected to the kind. However, connection type itself was not explicitly manipulated in the *zarpies* study, unlike in the *kevtas* studies and the pseudoword experiments, since they were more tailored towards testing concept acquisition in a narrative format.[23] Based on the general assumptions about overhypotheses, I argue that overhypotheses target PCs. This is due to the fact that the success of stipulating overhypotheses, in general, presumes that such a property is principally connected to the kind. This is because this theory relies on a mechanism that applies across various kinds, specifically because these kinds are connected by similarity of properties, via a superordinate kind.

[23] That being said, it would interesting for future research to design a new version of these picture book studies that separates principled and statistical connections as two distinct property types. This would allow for a mixed-design, with connection type manipulated within participants, and wording as a between participants variable.

Without coding properties into a binary system[24], this raises the wider question of whether any property that is mentioned alongside a novel animal is generally seen as more likely to be principled, by virtue of it being the only information explicitly mentioned[25]. Here, I will argue that these are two distinct phenomena. Leaving aside the contrast between PCs and SCs, I claim that if we continuously find that strong category-property links occur independently of subject form, a more likely explanation than linguistic variation is that they are caused by how our conceptual system links properties that are similar across entire animal categories, i.e. overhypotheses about locomotion, reproduction, diet, among others. Based on the post-hoc item analysis, I have argued the the properties that were rated to be most tightly linked to the category were 'eats tulips', 'buzzes when angry', 'wiggles its horn', and 'stands up on hind legs', which are expressions of diet, vocalisation, and (loco-)motion, respectively. Conversely, the least tightly linked properties were 'chases shadows', 'dances in circles', and 'bounces a ball on its back', all of which can be interpreted as idiosyncratic movement or skills.

Based on this, adopting overhypothesis as an explanation for the *zarpies* study results (and by extension as an explanation for the other novel animal kind acquisition studies presented in this thesis) leads to two implications: First, the linguistic restrictions that we observe between category and property type must in some form interact with more general conceptual expectations we hold about novel animal kinds. The results of the tasks will be modulated by more basic inference procedures of the conceptual system. This is similar to the observation that characteristics in isolation do not always, if at all, carry information as to how they are likely related to a kind. An example that illustrated the importance of combined category and property knowledge was that the property of *being red* can either be principally or statistically connected to the category, where there is an overhypothesis of animal colour, which does not apply to buildings (or the

[24]Or, indeed, an n-ary system, which includes striking properties and other groups of properties and how they are connected to a kind.

[25]This is apart from visual information in these picture book tasks, which, as I will discuss in more detail in the next chapter, should also be considered a principled property of an animal kind.

category artefacts), e.g. *Cherries are red* and *Barns are red*, respectively. However, this does not mean that the linguistic form cannot affect how we perceive novel properties to be linked at all. Instead, both overhypothesis formation and linguistic restrictions can be assumed to be balanced by the type of experimental task that participants are asked to carry out.

Second, and contrastingly, category-specific knowledge can grant certain properties a privileged status in induction and essentialism tasks. In this case, the knowledge about the relationship between animal categories and their means of locomotion, diet, etc. are part of the process of forming overhypotheses. These seem to have a privileged status in the human conceptual system. Thus, the results of the *zarpies* study address a cross-section of domain-general processes, linguistically-informed restrictions on generalisations, and category-specific expectations of essentialism.

I have now revised some of the assumptions made before carrying out the *zarpies* study. I have argued that the role of morphosyntax in concept acquisition can take on various roles. IS and DS subjects can be restrained by the type of property, such that they allow for principally connected properties but do not support merely statistically connected properties. This is most easily observed for linguistic restrictions of **existing** kinds, e.g. Lawler (1973)'s *A madrigal is polyphonic* vs *#A madrigal is popular.* Based on these limitations, the main purpose of these studies was an investigation into whether the reverse relationship also holds. In more concrete terms, these studies addressed the question of whether the linguistic form with which a **novel** kind is introduced could bias participants towards expecting a property to be present in novel instances of the kind.

The pseudoword studies in Chapter 3 and the *kevtas* studies in Chapter 4 have shown mixed results and demonstrated that various cues interact to arrive at a representation of a novel kind. Finally, the *zarpies* studies, which have taken a narrative approach, simulating how a child might learn about kinds through picture books and parental reading, have found no systematic differences in terms of how participants judged the link between categories and their properties, and neither in how essential they thought a property to be to a novel kind. All together, these studies have shown that while morphosyntactic cues may guide our conceptual system to expect certain properties to be more likely to be principally connected, other factors such as category knowledge can overwrite these assumptions.

The replication of the original *zarpies* study provided further support to the claim that generic, as opposed to specific language aids in generalising properties attributed of instances to novel instances. However, the effects predicted based on linguistic assumptions about the semantic representations of DEFINITENESS and NUMBER presented in Chapter 2 (and some of the experimental work presented in Chapters 3 and 4) were not evident from this experiment. Future research needs to address these methodological questions by simplifying the design to test category-property links in a more targeted manner and acknowledge the interaction between general cognitive processes in categorisation and those specific to linguistics cues. The boundary between properties that are likely principally connected and those that are likely statistically connected must be more clearly demarcated.

5.4 General discussion

This chapter addressed some fundamental questions about the acquisition of conceptual knowledge. Its main focus was to contribute to the literature on generalisations by testing experimentally hypotheses about the effect of building a more comprehensive mental model of a novel category, and doing so by using methodologies from both linguistics and developmental psychology.

By adopting and adapting the picture book methodology, I asked the more specific question of the role of linguistic cues in the acquisition of generic knowledge while relying on established literature. Previous studies lacked variation of linguistic form in their stimuli. As such, the results of these studies cannot fully account for the input to the conceptual system during childhood (and adulthood). This justified the replication of previous experimental work by Gelman et al. (2010) and extending it in precisely these terms, by including more linguistically informed stimuli.

The main findings from this study are that:

- Apart from the DS, all subject forms behaved similarly, including the non-generic, quantified subject *most*.
- Providing more context in this narrative style task, in addition to the category knowledge provided in the *kevtas* studies, leads to less distinct behaviour for each subject form.
- Due to the nature of the picture book study, the previously strictly controlled variable of connection type (principled vs statistical connections), was not a variable in this analysis. However, the section on overhypotheses illustrated that the connection between category type and property type needs to be taken seriously when investigating the question of how morphosyntactic cues and generic knowledge interact. This will be considered in future experimental design.
- The baseline for this study was *most* as opposed to the less conservative option *this*. While *most* should still behave differently from generic subjects that do not support merely statistically connected properties, the results of the *zarpies* study show no differences between the quantified baseline, the BP generic, and the IS generic subject condition.
- In this picture book study, the DS was often interpreted with specific reference, while the experiments in Chapters 3 and 4 demonstrated that they are interpreted generically like the BP and IS. This difference in interpretation of ambiguous subjects is important to keep in mind when contextualising the findings from more controlled and targeted experimental designs, such as match-to-samples studies, and the findings from more narrative style studies such as the corpus analysis in Chapter 2.

The inclusion of other generic subject types, both indefinite and definite singulars, was informed by the theoretical foundations laid out in Chapter 2 as well as the experimental findings in Chapter 3 and 4. In Chapter 2, I previewed some ideas about developmental input by analysing linguistic input in child-directed speech in English. In that chapter, I also conducted a corpus analysis, focusing on the presence, or rather absence of morphosyntactic variation in generic subjects for children. To recapitulate briefly, the findings support the hypothesis that children are mostly exposed to BP generic subjects, and IS and DS subjects with generic reference are virtually absent. DS subjects, which express kind reference in English, are particularly sparse. Similarly, the absence of kind-level predicates in the corpora (which normally co-occur with DS subjects but may also co-occur with BP subjects) points to a puzzle as to how these generic and kind-referring meanings are acquired for these ambiguous subjects. By contrast, BP generic sentences are abundant in child-directed speech. This is in line with the psychology literature, which argues that a lot of our knowledge is stored and expressed in generic form (e.g. between 1-1.5% of child-directed speech is generic phrases about artefacts, Gelman et al., 2005, Gelman et al., 1998), while relying heavily on the assumption the generics are best represented via BP subjects. This contrasting analysis provided a useful starting point for the experimental work in this chapter.

A partial answer to the question of acquisition of singular generic subjects was provided by the *kevtas* studies in Chapter 4. They have shown that even young children are able to interpret IS and DS subjects as generic. Simultaneously, it is essential to bear in mind that in these same studies, the responses to specific sentences, which cannot be interpreted as generic in adult speakers, indicated that children might generalise properties in a similar fashion as when presented with DS generic subjects. The older participants in these studies, aged 8-10 years, display almost adult-like behaviour. It is therefore yet to be determined whether young children have the same representation of generic reference for DP subjects as they do for BP subjects.

Both the corpus analysis and the findings from Experiments 1-5 (DEF-NUM), the French pseudoword study Experiment 6 (DEF-NUM), and Experiments 7-12 (world knowledge) helped inform the design and data analysis of the *zarpies* study. Its main purpose was to follow up on some of their partially contradictory findings. When comparing the findings of an experiment relying on conveying novel category knowledge in a narrative style (*zarpies*) to the more direct measure, but also less natural experiments of the *kevtas* studies in Section 4.2, the first difference to acknowledge is that the *kevtas* studies mainly relied on the *shape* bias. They worked under the premise that a *shape* similarity is always the preferred means of categorising animal kinds, unless certain characteristics are essential and pointed out through linguistic cues. With a one-off exposure, if the predicated *property* was interpreted as generalisable and essential, participants would choose the option dissimilar in *shape* over the the option without the *property*. However, without carefully controlling for the property type within the picture book and the test booklet, the results are comparable only to a certain extent. One might argue that the design of the *zarpies* test booklet conflates assumptions about inheritance or stability, which functioned as a proxy to essentialism, and could therefore render unclear response patterns due to unrelated considerations when answering the question booklets. The studies are comparable as they both introduce novel animal kinds and ask about novel members. However, in the *zarpies* booklet, there were other instances where novel members were specifically introduced as such and participants were asked to predict whether they possessed certain characteristics. This widened the range of assumptions that were tested, and at the same time left room to better understand other subtle, or even subconscious mechanism of inductions about novel categories. Thus, the benefit of the *zarpies* study lies in emulating a more natural way of acquiring new kinds, at least in the cultural context of the participants tested for this study, as well as the original studies run by Gelman et al. (2010). It can therefore provide a closer proxy to learning and organising properties in the context of more information, but linguistic in forms of the repeated sentence structure across the book, and varying visual stimuli.

The trade-off is that the results have complicated the picture by pointing towards more domain-general explanations rather than uncovering direct correlations between morphosyntax and conceptualisation. This can be attributed to the stimuli of the pseudoword studies evoking different cues and asking to carry out a different task. They should be taken as an attempt to emulate concept acquisition based on conveying knowledge through story-telling, unlike the more targeted tasks such as providing scores on a Likert-scale or matching a novel image to a target instance.

This is not to say that the lack of close correlation between subject form and scores for essentialism and category-property links did not shed light on concept acquisition from another angle. It opened up the discussion to consider alternative mechanisms in categorisation. This does not detract from the argument that morphosyntactic variation carries semantic information about the category-property links. Rather, it adds an additional layer to the mechanism under investigation by admitting that in the presence of visual information, additional category knowledge becomes available. For this study, this information pertains to the make-up of a land-dwelling animal. In this case, the judgements about category-property links for *zarpies* were the same whether the animal was introduced with a BP or IS, and differed for DS subjects. This shows that generic subjects, even under this broader task, can vary. By contrast, the scores were virtually the same for all subjects for essentialism questions.

Bringing together the conclusions across multiple studies in this thesis demonstrates that the different results of the *zarpies* study, the presence of category-specific knowledge (in this case the knowledge that the novel category is an animal kind) might overwrite subtle morphosyntactic differences in generic subjects. In a similar vein, in the absence of category-specific knowledge, linguistic differences have a bigger influence in our conceptualisation of the novel kind. Despite the availability of visual cues, the generalisations for each subject type are still affected by the task type the participants are asked to carry out. Ultimately, overhypothesis alone can also not account for the results and a joint consideration of domain general and specifically linguistic cues is necessary to better understand concept acquisition. In this way, the study enabled an investigation into some general patterns of specific properties, and introduced an alternative account by analysing the study under the lens of overhypotheses.

Moreover, by specifically looking at the different questions and methodologies between the experiments in this thesis, it begs the question of the effect of methodology on experimental results. This chapter leaves the puzzle of the role of morphosyntax in generic subjects unsolved. Instead, it has highlighted the role of task type and the form that stimuli may take, and illustrated that these variables jointly give rise to an interpretation of a novel kind and how its predicated properties are related to it.

To integrate this additional insight, the next chapter looks into potential task-dependent effects. It also investigates how they can be explained based on either knowledge-driven approaches (or theory-theory) as opposed to exposure-based (or exemplar- and prototype-driven) theories. Since there are systematic differences in the way psychologists and linguists run concept acquisition studies, as explored by replicating and extending well-attested paradigms throughout this thesis, these differences might deepen our understanding of the various results and findings. It presents the final set of studies of this thesis. Similar to the *kevtas* studies in Chapter 4, participants will be presented with novel instances of a kind, that are explicitly mentioned are a member of that kind (while other ones are explicitly described as **not** being a member). This will be framed within a novel version of the design: Those studies reverse the match-to-sample paradigm and thereby creates a set up that forces participants to choose one of various generic subject types. These studies will thereby also move away from the comprehension-focused design so far and more closely simulate a production-like environment.

It's perfectly possible that ***the dodo*** *would've eaten carbon monoxide and shat it out as cricket statistics. So where would you have been with* ***the dodo****? You would've loved* ***it****.*

— Nish Kumar: The Bugle, 12 May 2019

6
Methodology matters

Contents

6.1 Interim discussion

The overarching question this thesis seeks to answer is, "What is the role of language when learning about the world?" or, more specifically, "How can morphosyntactic cues in the subject influence the way we think about categories and the relation they have with their predicated properties?" The experimental research in this thesis has made a continued effort to explore this question from multiple angles. The studies in the previous three chapters have employed a variety of methodologies to explore the conditions under which the various structures that exist in English to refer to generic subjects are used. The results have indicated that the observable morphosyntactic differences, which represent underlying semantic differences, may affect expectations with regards to these category-property links. They have also shown that overall mechanisms of the conceptual system can override subtle morphosyntactic cues. This means that the effects were not reliable and left the interpretation of the results obfuscated.

The main question of this particular chapter then poses itself as, "Why do different types of studies give different results?" Or rather, "Why do some experiments yield results that are more closely in line with the predictions made based on linguistic intuitions?" I will identify theories that have given an account of the role of cues that affect categorisation and how they fare given the empirical evidence accrued in this thesis. I will investigate the variables that were manipulated between the *pseudoword* experiments in Chapter 3, the *kevtas* studies in Chapter 4, and the *zarpies* picture book study in Chapter 5.

- In Chapter 3, the results showed that morphosyntactic cues alone may raise expectations as to whether a category-property link is principled or statistical. However, these studies only provided limited and ambiguous evidence for this hypothesis. Moreover, they raised the issues of the amount of exposure to a property with a subject of a specific morphosyntactic form, and how this can affect the conceptualisation of category-property links.

- In Chapter 4, the results showed that generic interpretations are easily accessible for SINGULAR subjects in match-to-sample tasks. However, the role of morphosyntax might be diminished when simply pitting *shape* and *property* matches against each other, where participants can more easily rely on visual and auditory information. It also raised the question of the significance of using an artificial, one-off exposure paradigm as opposed to other studies which have used a naturalistic, narrative setting.

- Chapter 5 illustrated that general knowledge about category membership is a powerful tool to predict category-property links not only of novel instances, but also of novel instances of related kinds, if they fall under the same superordinate kind. These studies highlighted the powerful and possibly underestimated role of overhypotheses in tasks that aim to uncover the role of linguistic form (both generic subjects and *most*-quantified subjects) in learning paradigms that include category or world knowledge.

To recapitulate, the findings of a range of studies have not been conclusive. They differed in their methodologies, yet they all shared one aspect: They focused on linguistic comprehension. The differences in their results are hypothesised to be due to a combination of using novel methods as well as extending well-established types of experiments to ultimately bring together assumptions and results from the psychology and linguistics literature. Therefore, this chapter is dedicated to understanding the effect of employing multiple methodologies to answer the same question.

This means that after starting to investigate the effect of using one specific generic subject as opposed to another, we need to consider at least seven other variables that affected the results collected:

1. information regarding category membership (and hierarchically: superordinate, basic, subordinate kinds)
2. type of information available about novel kind (linguistic, visual, auditory, or any combination thereof)
3. contextual information (is the kind introduced as part of a narrative?)
4. amount of exposure (is the same or similar information made available repeatedly?)
5. task type (membership categorisation, agreement with statements, inheritance and stability judgements?)
6. comparison class and baseline (number of generic subjects, quantified and specific baselines)

7. linguistic mode (comprehension vs production of language)

Some of these variables were addressed experimentally and an improved understanding of their effect was gained through changes in the methodology from one experiment to the next. For example, the previous chapter concluded with an item analysis of the properties predicated of the *zarpies*, and provided an unexpected appreciation for the first variable, i.e. that independent information about category membership, which is inherited from superordinate kinds,[1] may have a stronger effect on the expectations of category-property links than morphosyntax alone. Another reason that these differences were not strong, and note that these points are not mutually exclusive, could lie in the methodology of Chapter 5's Experiment 13 (narrative). The tasks that the participants completed targeted their understanding of novel categories in a broad sense, including assumptions about inheritance and stability of the category's properties. These are common indicators of stability but not necessarily ideal identifiers of category-property links. While we accept that certain inherent behaviours are likely to be stable, it still allows participants to disagree with questions such as "Do you think this zarpie will always be able to *property*?", even if they believe the property to be principled.

The experiments presented in this thesis have shown that both the specific instructions and the amount of exposure affected participant behaviour. The wording of the experiments sometimes leads participants to form strategies that they adhere to throughout all experimental trials. This was supported by participant comments in the *kevtas* studies. Therefore, to improve the methodology and use a more appropriate measure for the research question that exclusively targets the category-property links, the design of the final two studies was changed to a version of a match-to-sample task. In Chapter 4, I showed that match-to-sample tasks reduce the variability of tasks on beliefs of inheritance of stability (unlike the tasks in Experiment 13 (narrative)). Given that people often rely on similarity

[1]Interestingly, this relates back to the differences I addressed in Section 2.5.1 regarding the different paraphrases used in Prasada & Dillingham (2006, 2009), contrasting the semantics of *by virtue of* and *one aspect of*. Both of these are used to target PCs, but they both yield different results as they give rise to distinct property-inheritance structures.

in shape to categorise novel instances, this is taken as the default option in this categorisation task. However, only if the property is seen as essential to the kind, do participants choose a novel instance that possesses the property, even if this instance is dissimilar in shape. Thus, while I have repeatedly shown that singular generic subjects are also easily generalised, like the BP, the observation that these three generic subjects do not have the same distribution, and fulfil distinct functions, remains to be supported better experimentally. This is the underlying motivation for the final two studies presented in this chapter, which tested whether forcing participants to choose one of the three generic subjects would unveil preferences based on the type of category-property link.

To specifically target some of the methodological assumptions that have informed the studies in this thesis so far, I will contextualise the empirical results of these studies by inspecting my assumptions and findings, and assess how they fare with regards to various theories' predictions. In Section 6.2, I begin by taking a step back to evaluate these findings from the perspective of exposure-based and theory-driven theories about how categories are formed. These are seen as the two main streams and are distinguished based on whether categories are formed based on exposure, which encompasses statistical models of learning and can be further subdivided into models of prototypes and exemplars (Section 6.2.1), or definitional knowledge, which relies on inferences formed upon evaluating mini-theories about concepts (Section 6.2.2). Based on that, I will then compare the effects of using verbal cues only with experiments where they are presented jointly with visual and/or auditory cues in Section 6.3.

The insights from considering the methodological differences of the studies also allow us to look at their similarities: All the experiments in this thesis have approached genericity from a comprehension point of view. By shifting closer towards simulating the production of generic sentences, the final set of studies provides a new angle on differences and preferences between choosing one generic subject type over another when expressing certain category-property links. This novel approach merges methodologies that have proven to be most insightful, such as predicating two distinct types of category-property links and using normed visual stimuli jointly with previously tested wording. It uncovers some biases in the match-to-sample tasks and shows the distinct nature of IS, DS, and BP generic subjects from another angle, providing further insight as to whether they each support distinct expectations. These two production-like studies are presented in Section 6.4. The general discussion in Section 6.5 ties together the results from this chapter with those of previous ones and relates them back to the theoretical assumptions that have guided the design of both these two experiments and earlier ones, once again looking at the interplay of the variables that have been manipulated throughout the various sets of studies in this thesis.

6.2 Statistical and theory-driven inferences

Let us begin by focusing on the number of exposures participants received about novel kind and about novel properties in the studies presented so far: The experiments in Chapter 3 trained participants on two types of properties that many instances of various kinds possessed. The experiments in Chapter 4, on the other hand, introduced a new kind with each trial. Each of these kinds was visually represented by three instances, one target image and two options, distinguished by similarity in *shape* and possession of the same *property*. By contrast, the experiments in Chapter 5 introduced one kind to the participants, with a total of 21 instances, 17 of which had properties associated with them. These studies were different as they evoked a picture book narrative which allowed participants to form richer representations of one kind in particular, as opposed to building a representation around a property based on co-reference with the morphosyntactic form of the sentence's subject. From a conceptual standpoint, these tasks require different levels and mechanisms of abstraction. While the first teaches you a lot about one kind, the second gives you one data point about one kind at each task during the study.

Rips (2001) and Leonard & Rips (2015) both discuss different kinds of conceptualisation and reasoning about kinds. The main distinction made in Rips (2001) is between reasoning based on so-called mini-theories (theory theories) and similarity-based theories (mainly exemplar theories, but this includes prototype theories as well).

These distinct routes that are proposed for the conceptual system lead to different mental representations that eventually affect the conceptualisation associated with the novel kind. Below, I provide an overview of these two theories of concepts. After exploring the different implications these types of conceptualising have on the cognitive system, I will present the design for a new set of studies that specifically try to disentangle these notions that were conflated in previous experiments.

6.2.1 Similarity-based theories: prototypes and exemplars

The central hallmark of similarity-based, or exposure-based theories is that they, unsurprisingly, rely on an individual being exposed to an instance of a kind. More fine-grained distinctions can be made based on the amount of training and exposure that is available to a learner. In this section, I assume that both prototype theories and exemplar theories are two versions of the same type of theory. Both of these categorisation mechanisms rely on the input of data points, although they are distinct in terms of how much input they require and how that information is processed. Theories that rely on training may include models that employ machine learning where algorithms extrapolate based on available data points and predict future outcomes. In order to make generalisations, these mechanisms need information about previous categorisation, which creates a feedback loop.

Note that within the literature focusing specifically on prototype and exposure-based models, even finer distinctions have been made. Most notably, there is the distinction between models that rely on probability, based on exposure to cues and cue validity, and prototype models proper, also described as distance-based models (which could also be understood as similarity-based models). These differences and models in general received much attention in the 1960s and 1970s (e.g. Vygotsky, 1962; Bruner et al., 1966; Rosch & Mervis, 1975; Rosch et al., 1976; Reed, 1972; Rips et al., 1973 among many others), and more debate than is relevant to the discussion at hand.[2] What matters for the overview and argument in this section is that the consensus was closer to treating family resemblance (or distance-based models) and cue validity (or probability models) as tightly interlinked processes of an overall joint process, rather than conflicting theories. In fact, I would hypothesise that they cannot be separated, but might receive different weights depending on the category type and the types of cues available to the conceptual system.

Prototype theories are based around the idea that there is one central member of a category around which other members of the category are arranged in increasing distance, typically represented by concentric circles. The members of a category that are deemed the most prototypical, or the ones that can be visualised as the central members in various concentric circles, are the ones that "bear the greatest family resemblance to other members of their own category and have the least overlap with other categories." (Rosch & Mervis, 1975, p. 599) In other words, the further out in these circles they are, the less central they are to the category by virtue of being less similar to the most central member of that category. Popular illustrations of this theory include categorisation of birds, illustrating natural kinds, as well as cups and mugs, illustrating artefacts. The well-known visual of so-called 'Birdiness rankings' is given in Figure 6.1, as in Aitchison (2012, p. 69).

This way of categorisation anticipates problems, most prominent of all the question of how to decide which member is the most representative of a kind. By the same token, the question arises as to which instance is the least representative, while still being a member of the kind. This last point is typically handled under an account of fuzzy sets or fuzzy boundaries (e.g. McCloskey & Glucksberg, 1978; Medin & Smith, 1981).

Now that I have laid out the basic features of prototype theories, the question remains as to how these categories are formed in the first place. Earlier, I discussed differences between basic and superordinate kinds, especially with regard to overhypothesis. Yet, reviewing family resemblances and the development of prototype theories uncovers another related aspect of category formation: cue

[2]For a series of experiments that specifically target differential predictions for exemplar and prototype theories, see Medin & Smith, 1981, although I will briefly discuss their approach below.

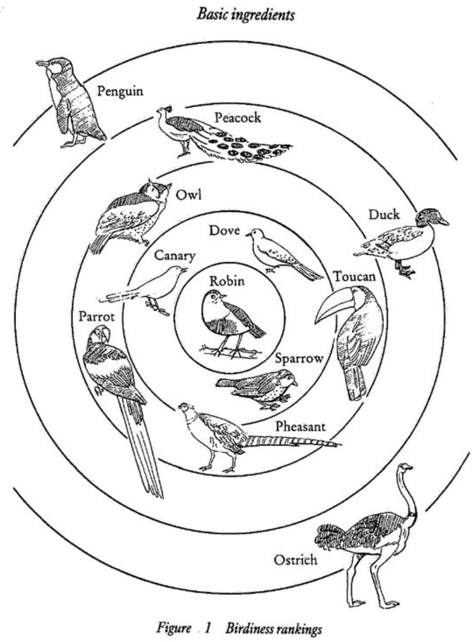


Figure 6.1: Aitchison (2012)'s birdiness rankings, illustrating prototype theory with bird category membership

validity. Rosch & Mervis (1975) argue that one difference between forming basic, superordinate, and subordinate categories lies in the richness of the available information, or the density of the information bundle, that any of these three levels conveys. For basic level kinds "in the domains of both man-made and biological objects, there occur information-rich bundles of attributes that form natural discontinuities" (1975, p. 586). By contrast, superordinate kinds share fewer attributes with one another relative to basic level kinds, as do subordinate kinds. This is why basic categories are conceptually so intuitive, they are "the categories for which the cue validity of attributes within categories is maximized: Superordinate categories have lower cue validity than basic [categories] because they have fewer common attributes within the category; subordinate categories have lower cue validity than basic [categories] because they share attributes with contrasting subordinate categories" (1975, p. 586f). The idea of cue validity is a recurrent one in Rosch' work, showing the relevance of probabilistic theories, which can express prototypicality as an independent measure. Note within this context that, while possibly not immediately obvious, the idea of an "information-rich bundle of attributes" lets us assume a theory in which there is not necessarily only one ideal, prototypical member around which the other instances of the kind or basic kinds are arranged (as pointed out in Murphy, 2002, p. 42).[3]

Rosch & Mervis make explicit that their theory is one of family resemblance, and not strictly of prototypes or cue validity, both of which they judge to be a more narrow class. In more specific terms, that means that the prevalence of a specific attribute is not enough but that the distribution needs to be considered as well. Precisely how that is implemented is empirically and theoretically borne out in the prototype literature, but the noteworthy component of her argument for the current discussion is that these attributes interact with other basic kinds that fall under the same superordinate kind. Rosch & Mervis summarise that the significance of distribution over mere frequency cannot be dismissed due to the high correlation that they found in their series of experiments between the ratings for distribution of attributes between categories themselves and the ratings of prototypicality. Critically, this correlation was not found for ratings of frequency and ratings of prototypicality. In addition to that, the nature of the kind is relevant insofar as it raises independent expectations of category-property links.

Having looked at exposure-based theories in more general terms, let us turn to the more specific issue of identifying members of a kind. Murphy (2002) argues that one problem with prototype models is that they do not differentiate between various types of information that we are exposed to when acquiring knowledge about new kinds, e.g. shape vs location. This warrants two follow up comments. First, referring back to the theory of overhypothesis, we understand that for animal kinds, certain characteristics are assumed to be consistent across instances of a kind and are united by second-order hypotheses. This theory can be extended to artefacts since they are designed for a specific purpose. Hence, their properties are assumed to serve teleological purposes, i.e. to serve the function that the

[3]This relates some of these deliberations to the developmental work I have discussed under different theoretical assumptions: Rosch & Mervis (1975, p. 587) mention that "basic objects were shown to be the first categorizations made by young children, and basic object names the level of abstraction at which objects are first named by children and usually named by adults", making reference to further (developmental) experimental work in Rosch et al. (1976).

artefact was built for.[4] Second, linguistic form tells the conceptual system how the properties are related to the concept, at least to some extent. Recall that the IS can denote properties that are prevalent but through accidental factors, such as habitat (as in the examples in the previous chapter, *A mole crab has muddy feet* or *A crocodile is wet*). Thus, learning about an artefact's shape vs location, these differences in linguistic form of the subject can further inform us as to how we are to conceptualise the associated property.

Interestingly, Fodor & Lepore (1996) relate these ideas back to issues of semantic compositionality. In short, they argue that in order to be productive, concepts must be compositional. As we have seen with the introduction of many novel concepts and hypothesising about hybrid concepts, they are highly productive. Fodor & Lepore continue that productive concepts cannot be prototypes. They take an extreme stance by which the two types of theories discussed in this section are irreconcilable, as "it is therefore as near to certain as anything in cognitive science ever gets that the structure of concepts is not statistical. Theories of categorization, concept acquisition, lexical meaning and the like, which assume the contrary simply don't work." (1996, p. 254) Their argument relies on the joint assumption that first, concepts are productive and they must therefore be compositional, and second, prototypes are not compositional and therefore cannot be the underlying structure of concepts. Over the course of this section, I will propose that a potential solution might include a joint theory of statistical theories and theory theories, which then can account for productive and compositional concepts. However, let us also bear in mind that the reason these theories are discussed at this stage is to evaluate how properties predicated of individual instances of kinds and properties predicated of kinds based on group averaging interact.

Much of this section has focused on prototype theories and family resemblances proper. As mentioned, finer distinction can be made depending on whether the focus is on distance-based or cue-based models. The main distinction is drawn

[4] A follow-up question regards whether the distinction between different kinds, i.e. animal kinds and artefacts, is innate or learned. This then, in turn, affects the origin of overhypotheses and how they can, if at all, support exposure-based theories.

between these prototype theories and exemplar theories. For the purpose of the discussion here, I take exemplar theories to be closely related to prototypes. They differ in that they represent individual instances, or a single exemplar, and can therefore not rely on the averaging of instances. However, they remain largely comparable since they are also built based on the perceptual features of encountered instances of a kind. Yet, without the averaging processes, they are not ranked in concentric circles as more or less representative members of a kind. This distinction is particularly pertinent when accounting for unexpected but normal members of a category, that should not be judged to be a less typical member based on sharing fewer characteristics with the 'ideal' category member. Moreover, exemplars could act as a shorthand to quick construction of categories, especially when there are only few past encounters to build on.

Medin & Smith (1981), with reference to Medin & Schaffer (1978), view exemplar theory more clearly as a contrasting alternative to prototype theories, where each novel member of a category provides information that can subsequently act as retrieval cues that form part of an associative network of the cues to retrieve category attributes of atypical instances. For them, the main difference lies in the role that any single exposure plays, and whether repeated exposure changes the statistical averaging procedures. Further, they raise the concern that in experiments, participants might form strategies early on in an experimental setting.

In fact, the way that single exposure and strategy formation interact was reflected in some of the comments provided by participants for the match-to-sample studies in Experiments 7-12 (world knowledge) in Chapter 4. While Medin & Smith (1981)'s experiments were different in terms of stimuli and hypothesis, they found a large effect based on the instructions that participants received, which is reflected both in my findings and the review of methodologies in this chapter.[5]

[5] As a methodological aside, consider that many pieces of evidence in favour of prototype theories come from a range of popular experimental tasks, e.g. picture-identification tasks, rating tasks, or agreement with statements, such as ratings on Likert scales. While these are easy to set up and replicate, as evidenced by the range of studies in this thesis, future research might enable a better understanding of the noticeable effects due to subtle manipulations such as the wording of the instructions and subsequently will allow us to put the findings into perspective and hopefully tease apart theoretical considerations in a meaningful way.

One way of reconciling these differences between largely similar theories is by acknowledging that these finer-grained differences exist, but that they do not posit underlying assumptions that are distinct to the extent that they cannot be explained by the same cognitive mechanism. While the information provided by an unusual instance of a category member, e.g. a flightless bird, raises the question of how they fit into a categorisation theory, ultimately both exemplar and prototype theories rely on (repeated) exposure to instances of a kind, or group members. Not only that, they both make assumptions about how good a fit any instance is based on the properties, visual attributes or other characteristics these instances have. As a side-note, this circumvents (although does not solve) the issue of whether being exposed to an instance, but not paying attention to it, constitutes being exposed to an exemplar.

Leaving aside how prototype and exemplar theories might interact, and in particular the effect of individual instances as opposed to groups and how the prevalence of their properties can be averaged across instances, I will ultimately turn to an issue that is more relevant to the methodological review in this chapter. Having seen the role of essentialism and overhypothesis in animal categories, it is necessary to narrow down broader considerations of exposure to instances of kinds and think specifically how different interpretations and experimental results might be explained by better understanding what participants knew about the tasks they were asked to carry out.

Kelemen & Bloom (1994) were among the first researchers in this area to seriously consider the role of pre-existing category knowledge (or domain-specific knowledge, as they call it) in categorisation tasks. Their starting point is the dominance of conceptual theories that highly emphasise the role of perceptual cues. Kelemen & Bloom (1994, p. 390) point out that under their so-called “rationalist” theory (see also Carey, 1985; Murphy, 1993; Bloom, 1999), perceptual similarity is not dismissed as irrelevant but is not in itself responsible for categorisation. Rather, similarity cues a person into expecting category membership with a similar-looking instance, since these instances likely share internal structure or function, again depending on category type. Kelemen & Bloom’s alternative relies on the idea that category-specific knowledge can never be taken out of categorisation tasks. They manipulated the way in which participants thought about the same type of visual stimuli by providing them with category knowledge that either classified these stimuli as belonging to natural kinds or to artefacts. They found that categorisation differed based on the category type that was provided to participants.

As we have seen in earlier chapters, this is knowledge that participants did not have in the pseudoword studies. This further shows how perceptual cues cannot be sufficient to account for distinct expectations of categories and how their properties are connected to them (in particular for Experiments 1-5 (DEF-NUM) in Chapter 3). In addition to these considerations, recall that at the beginning of Chapter 4, I mentioned a study run by Smith (1989) in which participants were asked to group shapes together, in particular circles. Adults grouped circles together that were identical in one of the properties, e.g. size or shape, and not those circles that were overall similar but never identical on any dimension. Interestingly, children did the opposite and matched overall similar circles. This is another reason to consider whether a *shape* match is the best baseline for match-to-sample studies, although it should be acknowledged that it is not necessarily an unjustified baseline.

Returning to Kelemen & Bloom (1994, p. 391), following up on Smith (1989), adult categorisation was tested once more, but with additional domain-specific knowledge provided to the participants. The relevant manipulation in this follow up study consisted of telling participants either that the circles were a natural kind, a condition in which they are hypothesised to group by colour, a proxy for pigmentation, or that they were machines, a condition in which they are hypothesised to group by size, a proxy for machine function. This prediction was borne out in their experiments. Curiously, the effect was stronger for animal kinds, which ties back to theories of higher essentialism for animals than artefacts.

Prasada & Hall (2019) present another argument, and caveats, on the role of exposure-based theories within the context of transformation stories. Their argument relies on perceptual similarity being a cue rather than a factor itself. These considerations go back to pitting and refining various approaches to prototype theories, such as Gelman & Markman (1987) or Keil (1989)'s transformations of e.g. porcupines into cacti, in which they retain their porcupine-essence and only assume an appearance that is cactus-like. These debates show that despite the fact that *shape* similarity can act as a baseline for categorisation tasks, and has done so successfully to provide insight into the conceptual system, there are broader discussions to be carried out with regard to what a more universal baseline looks like, either within or across category boundaries (if there is such a thing). It bears explicit reminding at this point that these findings work in favour of the stimuli type chosen for the match-to-sample studies in this thesis as the findings from earlier, independently motivated, research shows that animal kinds, even if represented as circles at microscopic level, default to categorisation by colour over size, which is a perceptually prevalent characteristic that the *shape* matches often displayed (but not always, and notably it was not overtly manipulated or considered in the stimuli design).

As we will see below, considerations addressing cue-retrieval based on exposure and feature density are not nearly as important in theory theories. Or, as Murphy (2002, p. 51) puts it with respect to seeing instances of llamas repeatedly over time: "The question is, then, when you have seen a few dozen llamas, are you forming a general description of llamas - as the prototype view says - or are you just getting a better idea of what llamas are like because you have more memories to draw on - as the exemplar view says?"

In general, for the experiments that I have presented so far, this provides further evidence that knowing an instance's category raises a priori expectations of how certain properties are linked to the category. However, these experiments were not designed to address further questions such as participants' strategies in the absence of any category knowledge, or a controlled manipulation of the linguistic cues that label novel instances. Yet, these were precisely the types of considerations that I have manipulated in the studies presented in the previous chapters and which have influenced participants' categorisation patterns as well.

What this means for a more general account of categorisation and the human conceptual system is that these strategies likely interact and that the human mind likely makes use of available cues as much as possible. The more specific knowledge is available, for instance in the form of category knowledge, or in the form of combined verbal, visual, and auditory cues, the more previous experience is drawn upon. On the other hand, in the absence of this information, more emphasis is put on the cues that are available, potentially beyond the extent that they would normally be emphasised. One possible consequence of the increased attention to linguistic cues is that their role is over-emphasised and thus influences category-property links in an unusual manner. As soon as additional types of cues become available, the significance of these cues is scaled back.

With this in mind, it is possible to assume that the support for prototype theories is heavily dependent on the types of tasks that have been employed to explore the nature of categorisation. By contrast, what the other main theory, the "theory" theory (sometimes called "rationalist" or "knowledge-driven" theory) relies on, is exactly what some of the experiments addressing the formation of prototypes have found: Categorisation draws upon pre-existing knowledge about how categories are structured, whenever possible. They do not solely rely on visual or linguistic input to form categories, but rather they implement visual and linguistic information into a blueprint of a (superordinate) kind. In other words, "categories emerge from the interaction between background knowledge and perceptual information; they are not abstractions from perceptual experience." (Kelemen & Bloom, 1994, p. 393)

I have discussed in detail the effect that wording and even the type of label has on participants' categorisation choices in the context of Experiments 7-12 (world knowledge). I would therefore be careful to draw broad conclusions based on these findings, seeing that the type of instruction was not manipulated in the same way that the instructions were for the experiments presented here (and recall the differences between Experiment 7 (world knowledge) compared to Experiments 8 and 9). The alternative approach to categorisation of relying on mini-theories of categories will be explored in more detail in Section 6.2.2 below, and will complement the ideas presented in this section. Taken together, they motivate the two studies presented in Section 6.4, which were designed to shed light on the effect of methodology and task type in the categorisation tasks presented in this thesis.

6.2.2 Theory theory

Having evaluated exposure-based thories, it looks as if it is inevitable that we draw upon prior knowledge in categorisation whenever it is available. In addition to telling the conceptual system what an instance of a kind is like, it can also be useful in situations where kinds are entirely new. This is mainly due to the way the conceptual system draws analogies (Murphy, 2002, p. 191ff). Therefore, we need not discredit the findings in empirical domains, where knowledge and categorisation are influenced by exposure and typicality effects. Instead, these might work in tandem, providing the conceptual system with any type of knowledge available, which in turn is constantly re-evaluated and integrated into our pre-existing conceptual understanding. Murphy (2002) summarises that there are "two different ways that knowledge might affect categorizations: (1) indirectly, by changing the concept representation during learning (i.e., before categorization); or (2) directly, by activating the knowledge during the category judgment. **It is likely that both effects occur at least some of the time.**" (p. 192, emphasis added)

Before we turn to how these seemingly distinct modes of category formation might interact, let us first look at the assumptions of "theory theory" (or "the knowledge approach"). Some of the arguments in favour of these theories have been developed independently of prototype theories, relying on the idea that we keep, and update, definitions of any category (as argued in e.g. Smith & Medin, 1981). The main appeal for theory theories is that they rest on the assumption that learning does not take place in isolation. Unlike prototype or exemplar theories, being exposed to an instance does not mean that this directly and exclusively affects a person's concept of that kind. It takes a more integrated approach, where encountering any type of animal evokes background knowledge about natural kinds and intermediate kind levels, e.g. knowledge about mammals or reptilians (Murphy, 2002, p. 60). Murphy continues that this relation is reciprocal, something that neither of the exposure- or distance-based theories could account for: "a new concept can also effect a change in our general knowledge. Thus, if you learn a surprising fact about a new kind of animal, this could change what you thought about biology in general (e.g., if you learn that snails are hermaphrodites, your knowledge about sexual reproduction in general could be affected." This allows for an explanation of the behaviour displayed by children (and experimental participants in the previous chapters) who quickly form categories for novel animal kinds, based on what they already know about other animal kinds.

Does this mean that we need to discard prototype and exemplar theories entirely, in favour of theory theories? One issue with proposing underlying mini-theories for concepts is that the nature of these theories on which concepts are proposed to be based is still debated. Moreover, recent experimental work, such as Leonard & Rips (2015), shows that people might be able to entertain multiple theories for a concept. Leonard & Rips conducted a series of experiments that asked participants to judge category membership and continuity of artefacts and natural kinds throughout various transformation scenarios. Particularly interesting is the finding that participants explicitly state that there are multiple ways of counting artefacts, especially if they are partly (dis-)assembled or only temporarily disassembled. In other words, they choose response options such as "(Both are correct.) There is more than one correct way of counting tables."

Leonard & Rips clearly discount a psychosortalist theory of concepts (as proposed by e.g. Rhemtulla & Xu, 2007, and more loosely related Prasada et al., 2002), and instead propose an "anti-psychosortalist" theories along the lines of Rips et al. (2006) and Blok et al. (2005). Leonard & Rips (2015, p. 90) describe psychosortalism as the theory of relying on mental representations to individuate and categorise instances into categories based on sortal terms (e.g. PERSON). An anti-psychosortalist approach, by contrast, does not rely on mental representations to trace object identity (p. 100) The main conclusions they draw is that judgements about individuals do not necessarily rely on accessing basic kinds or superordinate kinds (Leonard &

Rips, 2015, p. 100), and that transformations across basic level kinds are easily accepted, as exemplified by sentences such as *This school used to be a post office*. While their replies and refutes to psychosortalism shed some light on conceptual theory based on mini-theories, they do not provide the reader with a sufficiently fleshed out theory about the structure of concepts to the same level of detail that the prototype theories above did. While this might initially sound unsatisfactory, and potentially presents a weaker argument, it could provide insight into the structure of the mind that is more in line with independently found claims. Rips (2001) draws on independent support from neuroimaging studies (Goel et al., 1997; Osherson et al., 1998), showing that different brain areas are engaged in argument evaluation, depending on the inductive and deductive tasks, where inductively strong reasoning could be taken as a proxy for a knowledge-driven or theory theory of concepts. Within his own area of research, Rips' assumption is then unproblematic, as it aligns with the idea that inductive and deductive processes provide separate components to argument evaluation. Combining these findings with the distinct findings for assumptions about category-property links, they can provide independent support for the idea that category-specific knowledge is activated depending on the type of information that is readily available, as well as the manner in which it is presented (e.g. visually or exclusively linguistically). Ultimately, Rips (2001) too discusses how a unitary approach to reasoning seems unlikely given the types of inductive and deductive inferences we make. Instead, our minds seem to be able to rely on various modes to evaluate arguments, organised along a spectrum, to choose an appropriate mode of reasoning given the data and context. When participants were asked to evaluate arguments based on whether they were deductively or inductively sound ("valid" and "plausible", respectively), the responses differed as a function of whether the argument was deductively (in)correct and causally (in)consistent. As these judgements are distinct, a unitary view of reasoning seems unlikely. The advantage of such a dual model of reasoning is that it allows us to evaluate arguments based on the information provided and the purpose of the evaluation.

Earlier work has already proposed the insufficiency of either a theory-based or an exposure-based account. In particular, Murphy & Medin (1985) provide a detailed overview of the advantages and disadvantages of both of these theories, summarised in their Table 1 (p. 298) (repeated here as Table 6.1) to illustrate how these two theories can not only interact but also compensate for their respective shortcomings. Particularly relevant to the discussion at hand are the rows “concept representation”, “category definition”, “weighting of attributes”, and “conceptual development”.

Table 6.1: Murphy & Medin (1985)’s Comparison of Two Approaches to Concepts

Aspect of conceptual theory	**Similarity-based approach**	**Theory-based approach**
Concept representation	Similarity structure, attribute lists, correlated attributes.	Correlated attributes plus underlying principles that determine which correlations are noticed.
Category definition	Various similarity metrics, summation of attributes.	An explanatory principle common to category members.
Units of analysis	Attributes	Attributes plus explicitly represented relations of attributes and concepts.
Categorization basis	Attribute matching.	Matching plus inferential processes supplied by underlying principles.
Weighting of attributes	Cue validity, salience.	Determined in part by importance in the underlying principles.
Interconceptual structure	Hierarchy based on shared attributes.	Network formed by causal and explanatory links, as well as sharing of properties picked out as relevant.
Conceptual development	Feature accretion.	Changing organization and explanations of concepts as a result of world knowledge.

The previous sections have shown that the idea that there are multiple routes to conceptualisation is increasingly plausible and has been brought forward by a range of researchers. This field of enquiry has developed over at least the last 70 years, starting with seminal works such as Wittgenstein (1953, published posthumously), picked up again by Rosch & Mervis (1975), and is still debated in much detail as recent as Leonard & Rips (2015). If we acknowledge that at least some of the differing results are due to inconsistencies and general changes in the methodology (as well as stimuli), which could affect the kind of conceptualisation that participants employ, then changes to the task type while still relying on the same stimuli should also provide different results. This relates back to findings in Murphy & Medin (1985), who agree that there is likely an influence of causal explanations when acquiring statistical knowledge, and vice-versa real world co-occurrence can enhance and constrain the causal theories people posit. They thereby provide further evidence that people likely learn both based on causal explanations, or theory theories, and statistical co-occurrence, or exposure-based theories.

Another such change of methodology and stimuli was raised in McNorgan et al. (2007)'s general discussion. After stipulating different knowledge types based on statistical co-occurrence as opposed to causal relations, and considering how the type of knowledge influences the task type, they conclude that "small numbers of learning trials promote causal reasoning, large numbers of trials promote statistical learning. Most category learning experiments use relatively few learning experiences. In contrast, undergraduate students presumably have vast experience with the sorts of object concepts typically used in studies such as the present one." (McNorgan et al., 2007, p. 427, with reference to studies by Tangen & Allan, 2004). These same experiments often fail to demonstrate the influence of statistical co-occurrences, while only providing participants with little exposure (and arguably for the good reason of practicality and feasibility). Recall that the amount of exposure was a variable that I manipulated in the *pseudoword* studies. They showed that for this type of study, the amount of training had to be 'just right', where both too little and too much training did not render findings that were as distinct. This is because they likely introduced other constraints, such as cognitive resources or fatigue. Looking into the amount of training in more detail, one might change the design to provide participants with multiple exposure periods which might enhance the results (although see e.g. Gelman et al., 2010, which show that these types of training studies render varying success).

This reflection on methodology and theories illustrated distinct routes to conceptualisation and have motivated the studies in Section 6.4 below. Before moving on to this final set of studies, let us briefly look at the effect that the type of stimuli might have on participants' decisions of ascribing and generalising properties of instances of novel animals to the entire kind.

6.3 Comparing verbal, visual, and auditory cues

Three of the pseudowords studies in Section 3.3 and the picture book studies in Section 5.2.3 rendered results that emphasised the role of morphosyntax to different degrees. One question that has remained largely unaddressed concerns the main contributing cues available to participants in these experiments. Most notably, the *pseudoword* studies used linguistic cues only, whereas the *zarpies* studies employed verbal and visual cues in conjunction. These data from this range of studies, all of which support the claim that generic language in general helps generalise novel instances, show only limited support for the hypothesised distinct patterns for BP vs IS vs DS subjects.

Let us begin by comparing the *pseudoword* studies to the *kevtas* and *zarpies* studies, both of which included visual stimuli in addition to the linguistic input. The main takeaway point is that the effect of morphosyntactically dependent differences was varied throughout. Of the pseudoword studies, only Experiment 1 (DEF-NUM), Experiment 3 (DEF-NUM), and Experiment 4 (DEF-NUM) showed an interaction between wording and property type. Of the *kevtas* studies, only the final three adult studies showed this interaction, when including the order in which the property types were presented. The easiest conclusion would be to hypothesise that by including additional visual information, the role of morphosyntax is less pronounced. This might be due to the visual salience of colourful pictures of novel animals, combined with the formulaic pattern of presenting these animals which allows participants to skim over the linguistic form and simply rely on the content words in the predicate to make their decisions about which animals to group together. Seeing that Medin & Smith (1981) argued that participants easily adopt strategies in categorisation studies, and that these are heavily influenced by the instructions, this is a possible explanation.

Further support for the overriding effect affected by visual stimuli comes, perhaps surprisingly, from the *pseudoword* studies themselves. In the discussion of the studies, I raised the point that participants often commented that they did not understand what the point was and that they felt confused. Yet, these studies repeatedly showed that the morphosyntax of the subject raises expectations as to the link that the category has with its property. I concluded that these links are often not part of any conscious categorisation process and are instead automatic and subtle in linguistic processing. If this is the case, then we need not stipulate that recalling linguistic form is necessary for the linguistic pieces to affect our conceptualisation of their characteristics and result in distinct behavioural patterns. What is minimally necessary, however, is that these pieces are processed at some point. If, instead, the participant is able to primarily base their categorisation decisions on visual input, the linguistic pieces seem to be demoted in relevance.

I caveat this simple explanation by pointing out that the baseline condition in the *kevtas* studies consistently showed that participants are at chance level between choosing the *shape* and the *property* match. Moreover, by removing the label that introduces the target instance (from Experiment 9 (world knowledge) onwards), and thereby reducing the different labels, all of which can be interpreted as generic subjects, the IS condition was more in line with the hypothesis. This could indicate that some morphosyntax is less affected by the presence of other information, such as visual information. While this has not been explicitly tested, these cross-study comparisons may indicate how these factors are conflated. Changing the label and norming the stimuli helped refine the differences and show that at a coarse level the BP, IS, and DS are distinct. Mainly, however, these experiments showed that all three subjects are interpreted generically as opposed to their episodic reference which is often expected, particularly of the DS. This raises the question of whether the task type itself smoothed over the different interpretations for this study, as it might have been perceived as irrelevant, or whether the visual stimuli were the reason these three generic subjects were treated more similarly than in studies with linguistic-only cues.

Finally, we saw that including auditory information did not have a significant effect on the results of the *kevtas* studies. In the studies with adult participants, the results were slightly less pronounced than they were for the non-auditory version. This might be due to the randomised presentation of the stimuli rather than the blocked presentation of PCs and SCs. For the child studies, all three generic subjects were highly accepted, and PCs were rated higher than SCs across the board. The difference between PC and SC scores was largest for IS and BP subjects. Additionally, including auditory cues seemed to enhance the perceived educational or instructional tone of the studies, as the high scores, indicating a generic interpretation, were relatively high for DS subjects as well as the control group.[6] This rests on the assumption that an educational tone in the presentation of properties of novel kinds seems to indicate that these properties are central and therefore generalisable, highlighting the importance of the category-property links. What has not yet been tested is a combination of non-visual linguistic-only phrases that are presented with auditory information. The reason is that the inclusion of auditory information was to test children that cannot read yet - and the inclusion of pre-recorded and manipulated stimuli was to ensure that any differences in interpretation were not induced by experimenter intonation. Seeing that auditory-assisted reading is an important field of research, e.g. in second-language acquisition, this might be a separate way to test generic interpretations in both children as well as second-language learners.

The other experiment in this thesis that included visual stimuli was the *zarpies* study in Section 5.2. Its main aim was to provide greater insight into the relevance of contextual learning and the acquisition of a novel animal kind in a story-telling paradigm. While this design did not enable us to tease apart the underlying semantic and morphosyntactic differences in behavioural studies to the extent that the more targeted previous studies did, it raised the issue of adequately valuing the effect of knowledge about superordinate kinds. I hypothesised behavioural differences based on the morphosyntax of the subject, similar to the ones in the *pseudoword*

[6]This lead to changing the auditory stimuli once more to include a middle sentence that breaks up the repetition of the generic (or baseline) sentence for follow-up studies.

and *kevtas* studies. These were not found due to the considerable overlap between the answers of the BP and IS conditions, and some of the DS responses, which did not show any interpretational differences between the three generic subjects. However, this study once again confirmed that interpretations of singular generic subjects are easily accessible, even when presented in a picture book. Interestingly, the exclusive use of DS subjects led to participants either answering questions about category-property links and essentialism of properties in line with BP (and IS) subjects, or below the rates found for the *most* quantifier version of the book, showing that only using DS subjects is unusual.[7]

This discussion leaves unanswered the question of why visual cues are more powerful than exclusively linguistic cues. It is unclear whether the variation in the morphosyntax is taken more seriously when no other cues are available to a participant. What was observed is that the differences between the generic subjects are more emphasised when they are the only cue available compared to when (perceived) disambiguation can be provided by additional cues. Exposure-based theorists might argue that the characteristics about which participants learned built up a representation of the novel animal, whether that be an exemplar or a prototype, that highlights both *shape* and *property* preferences. Since the prototype around which the *zarpies* were built is never seen, the category boundaries might have been proposed to be fuzzy enough to propose essentialism that extends to many novel instances. A *Generics-as-Default* theorist might propose that due to the cognitive default of generics, combined with the high essentialism generally found in animal categories (thereby also including overhypothesis theories that have been hypothesised beyond animal categories and of bags of marbles), participants did not have a reason not to generalise based on any subject that might be generic. In the absence of category knowledge, visual cues, and *shape* information, participants had to "make do" with the minimal information they were given, thus only relying

[7]The results were more in line with the specific demonstrative *this* used in the original studies. For more detail, refer back to the discussion about the high ambiguity for DS subjects in picture books, as well as an analysis of the distribution of generic subjects in existing picture books about animal kinds and about artefacts that elaborates on this observation in Section 2.8.2.

on what they had acquired about the semantic differences between generic subjects, even if that information was not explicit (just as overhypothesis or cue-based theories would not be explicit to a participant or, in fact, anyone categorising instances). The studies can at best lead to this educated guess, but should provide enough context and experimental findings to pursue in more detail the interaction of stimuli type and category-building theory.

This leads to two main takeaway points from the methodology and stimuli review: First, having included singular generic subjects in experimental paradigms has taught us that generic interpretations are much more readily accessible than has long been assumed in the literature, regardless of stimuli type. Consequently, relying on the BP due to its ubiquitous nature and ability to generalise both PCs and SCs (as well as striking properties) is (a) not necessary, since both IS and DS subjects, despite being less prominent in everyday language, allow for generalisations in experimental settings, and (b) a shortcoming as, just because it is prominent, does not mean that its semantics are fully understood. In many ways, the BPs ability to generalise most properties is more elusive than testing a singular generic subject for which the semantics seem clearer. This is, of course, inclusive of the general issue of variable truth-conditions, which concerns BPs more than it does the more restricted IS and DS.

Second, the role that contextual information plays has likely been underestimated in previous studies, as well as introspective judgements of minimal pairs. Among those who have considered contextual factors, Lazaridou-Chatzigoga et al. (2019) discusses the role of domain-restriction of generics and Sterken (2015b), among other philosophers, addresses context-sensitivity of generics. It is essential to take these findings seriously as they show that many of the longer standing assumptions about the truth-conditions of generics might not hold if put to the test in the way they have been by linguists and psychologists. Even simple manipulations such as the wording of the label in a match-to-sample task, or the extent to which the generic phrases sound instructional, have effects on the acceptability of property types with certain subject forms.

To address the points made in this section, the final set of studies in this thesis will combine visual and linguistic cues more explicitly by asking participants to think more actively about the role of the linguistic form. Since the role of morphosyntax often seems conflated by being interlinked with other cues, such as visual and auditory stimuli, the next section reverses the match-to-sample paradigm. This way, the wording manipulation is not between participants, but the dependent variable that is evaluated within participants. This is the final set of studies presented in this thesis, designed to follow up on the role of methodology, which has been the focus of this chapter so far.

6.4 Asking to give generic or specific instructions

Chapters 3 - 5 have shown that it is easy to access generic interpretations, and to do so with any generic subject. More so, all generic subjects allow us to generalise properties and the experiments have consistently shown a tendency towards expecting a property to be more likely principally than statistically connected to a kind. These category-property links can be modulated when designing stimuli with carefully attributed properties. Ultimately, this should provide access to the more fine-grained differences pointed out in linguistic intuitions, which motivated the studies in Chapter 3, and could then be enhanced in other paradigms that were initially developed in psychology studies (see Chapter 4 - 5 in particular).

However, these differences can be subtle when tested in behavioural paradigms designed to target generic language as opposed to non-generic language, and not to target differences within generic subjects. This raises the question of whether the differences hypothesised strongly in the linguistics literature may become more clear-cut when participants do not rely on various types of generic wording to make decisions about whether another animal belongs to the same category when presented with two novel samples to match to the initial target. Instead, it might be possible to reverse this paradigm and present participants with information about which of the two samples is indeed another member of the novel animal category.

Then, they could be tasked with choosing the wording that they judge to be the most appropriate to convey the visually presented categorisation.

The strongest connection between all of the previous experiments in this thesis is that they are all similar in terms of investigating the **comprehension** of generic subjects. However, teasing apart the subtle differences between generic subjects can additionally be done by adopting a production-like methodology. One could provide participants with multiple options between which they may choose, and thereby move away from learning and acquiring novel concepts. To explore this idea, the studies in this section presuppose that a participant already possesses all relevant information about what a member of a novel kind looks like. This paradigm thereby falls more closely in the realm of **production** studies. Even though it offers participants limited options in a forced multiple-choice setting, it could reveal production preferences for PCs and SCs, and possibly the interactions that were observed in the comprehension studies in the previous chapters.

Seeing that the *zarpies* studies gave insight into second-order generalisations via overhypotheses, but unfortunately treated all generic subjects the same (when interpreted in their generic function), this study returns to separating category-property links, and thereby PCs and SCs, in an extremely targeted manner. These studies rely on the distinctions made for the *kevtas* studies (Experiments 7-12 (world knowledge)) to provide a consistent baseline with which to compare the results.

Finally, there were considerations in earlier chapters with regards to *Generics-as-Default* and the role of reaction times. For these studies, this will not be the main focus, and there are certainly more precise experimental techniques to collect reaction times and to time-lock them to the exposure to one specific stimuli. But bearing in mind their potential significance, they were once again collected so that they can be used as a first pass on which to base future studies that are more interested in the on-line unfolding of generic interpretations. Here, they will serve as an estimate to uncover some general tendencies between different generic subjects. This is similar to the *pseudoword* studies in Chapter 3, where they have been discussed as a potentially interesting variable. Again, they should be seen as an approximation of processing times rather than the main measures of these behavioural studies.

To summarise, the two studies in this chapter will address some of the variables enumerated in Section 6.1 above: First, what is the effect of using a different linguistic mode? The studies here are the first ones in this thesis to emulate production more than comprehension in a learning paradigm after a training phase. Second, task type is manipulated as this study offers multiple choice options for the responses. Previous studies have either used a Likert-scale to express the level of agreement with a statement (as in the *pseudoword* studies), offered participants to make a binary choice between a *shape* and a *property* match (as in the *kevtas* studies), or given them an entire test booklet with various tasks to complete (as in the *zarpies* studies). Apart from that, the variables from the previous *kevtas* studies are held constant to ensure that any differences are not due to new visual or linguistic stimuli, properties or contextual information.

6.4.1 Experiment 14 (production): Reverse paradigm

The two studies presented in this section return to the underlying idea of the match-to-sample task in Experiments 7-12 (world knowledge). In those experiments, participants were shown triplets of images of novel animals. The target image is accompanied by a generic (or specific control) sentence based on which the participant chooses one of two images which is "another one" of the same kind. One image is similar in *shape* and the other displays the target image's *property.* This paradigm is based on the assumption that people have a *shape* bias when categorising novel (animal) kinds. However, if participants generalise the predicated property beyond the target image's instance, this can override the *shape* bias and push participants towards matching the *property* image with the target one.

I adapted this methodology by showing these types of triplets of novel animal pictures, only providing a label and additionally circling either the *shape* or *property* match. Participants were told that the circled animal is a member of the new animal kind and that the other one is not. They were then asked to choose the sentence that best describes the relationship between the two animals, out of four sentences which differed only with regard to the subject form. The sentences had BP, IS, DS, and specific (*this*, control) subjects, displayed as multiple-choice options. I systematically varied whether the predicated properties were principally (PC) or statistically (SC) connected to the animal, e.g. *Bactras have three legs* or *Bactras tap on bottles.* This resulted in a 2 (property type: principled vs statistical) × 2 (match type: *shape* vs *property*) within-participant design, such that each participant saw half of the animals with PCs and half with SCs. Based on their semantic representations, I hypothesised that IS subjects are preferred for *property* matches and PCs. I expected similar patterns for DS subjects, but at a lower magnitude due to their more restricted use. Finally, I expected that BP subjects would be equally acceptable for PCs and SCs, but be used when the images represented a *property* match. For *shape* matches, I expected participants to choose the control condition.

I caveat these hypotheses by pointing out that they require metalinguistic awareness of the participants that they might not have (or if they do, they do not normally need to tap into it). Thus, they might continue to rely on their default mechanism for generalisations and employ BP subjects at a higher rate.[8] While this concern should be taken seriously, especially when introducing a new experimental method to gather empirical data, I would like to point out observations made by philosophers concerned with metalinguistic negotiations, which require metalinguistic awareness, such as Plunkett (2015); Plunkett & Sundell (2019); Sterken (2015a,c, 2020). Their main argument is that metalinguistic negotiations, i.e. negotiations about the usage of words are ubiquitous in everyday language. Especially conversations about gradable, but also non-gradable, adjectives are subject to persistent debates. Plunkett (2015); Plunkett & Sundell (2019) and Sterken (2020) provide examples of everyday interactions where interlocutors mediate their expectations of category meaning and normative expectations thereof (1) (first

[8] One way to prevent this is to exclude the BP and only provide IS and DS generic subjects for multiple-choice responses that express PCs. I will explain this alternative design in more detail below.

discussed in Ludlow, 2008) or subjective standards related to subkinds (2), as well as their expectations of category-property links (3)-(4), thereby easily drawing on their metalinguistic awareness and how language should be used (all examples based on Sterken (2020)'s presentation).[9]

(1) Talking about the race horse Secretariat.
A: Secretariat is an athlete.
B: No, Secretariat is not an athlete.

(2) Talking about the chilli served for dinner.
A: A red hot chilli is spicy.
B: No, a Habanero chilli is spicy.

(3) A: Madrigals are polyphonic.
B: Yes, though madrigals can be monophonic.

(4) A: A frozen dessert is best served immediately.
B: Some people like melted ice cream.

Based on these examples, we can assume that everyday interactions do not only presuppose a certain level of metalinguistic awareness, but that people effortlessly make use of their knowledge. This can be employed to establish common ground in a conversation as well as to weaken and strengthen the (modal) content of a

[9]To provide a more comprehensive picture of some of the implication of such theories, note that one consequence is that, due to the widely differing modal content of generic statements, there is no reason to assume only one type of quantificational force for generics. Based on their context-sensitivity (which has been independently asserted by Lazaridou-Chatzigoga et al., 2019), generics can provide such a variety of meanings, and those may be strengthened or weakened by the interlocutors as necessary. This means that there is no good reason to assert different semantics for different generic subjects. Of course, this is not a desirable outcome for the arguments put forth in this thesis. One test that might show that context-sensitivity and differential semantics based on distinct morphosyntax may co-exist would be to test the different generic subjects in different contexts, to see if they still display unique behaviour even if guided towards or away from one specific interpretation. I will briefly put this idea forward as an idea for a continuation of the reverse match-to-sample studies in Section 6.4.2.5 below. Another reason not to believe that there are no differences expressed in the generic subjects and the type of category-property links they evoke is that assuming that all generic subjects can express the same modal content, why would every language maintain linguistically distinct forms? Not only does that seem cognitively unresourceful, it also leaves us to explain minimal pairs that rely on these differences, such as *We can't change assumptions about what a leader looks like unless we change what leaders look like* (Filipovic, 2016).

generic expression. Plunkett & Sundell (2019, p. 1) admit that the extent to which interlocutors are aware of the “disagreement (...) about how they should use language in the context at hand” varies and state that it might result in a cost to debate the theoretical underpinnings of a topic that is not at-issue. However, as participants in this experiment are already engaged in non-naturalistic linguistic interactions, I will rely on the assumption that they have these skills at their disposal, and that they do regularly engage in using these strategies to negotiate metalinguistic issues.

In summary, this type of experiment presents participants with the triplets of novel animals that were used in the *kevtas* studies in Chapter 4, however without any of the generic or specific information that was available to participants when they were initially presented with the triplets. Instead, the task for participants was to choose one of four sentences that best describes the scene they saw, i.e. the one that best represents the relationship between two of the stimuli, the target image and the match.

6.4.1.1 Participants

161 monolingual, native speakers of American English were recruited via Prolific. They were randomly assigned one of 4 conditions (n $\approx$ 40 per condition). They were aged 18-50 and had no (self-reported) language-related disorders. They gave informed consent and received financial compensation (£0.50) for their participation, which took approximately 5-10 minutes.

6.4.1.2 Items

The stimuli were based on the *kevtas* experiments in Chapter 4. Participants were presented with the same triplets as in the *kevtas* studies presented in Sections 4.2.1 through 4.2.3. However, for each triplet, one of the two sample images, presented below the target stimulus, was circled. Further, instead of seeing a description of the novel animals in form of a generic sentence (or a specific control sentence), the task for the participants was to answer a multiple-choice question. They were instructed that the target image was a novel animal, e.g. a ‘kevta’, and that the circled sample image was also that same animal. Crucially, they were told that the animal that was not circled, was **not** the same type of animal. Based on this information, they had to choose one of four sentences that best described this relationship. The full instructions given on each trial were: "Which of the following four options would you use to describe the first picture to someone, so that they would also think the circled picture is a bactra, and not the other one?" An example of a full trial is shown in Figure 6.2.

The conditions were randomly assigned based on a Latin-Square design. The 8 novel animals were split up into two sets as in the previous *kevtas*-studies. Set A included *bactras*, *feps*, *screds*, and *febbits* and Set B included *vorzyds*, *plogs*, *bants*, and *kevtas*. Within each set, the connection type that was presented in the answer choices of the multiple-choice options was either principled or statistical. However, the target-sample similarity, i.e. whether the *shape* or the *property* match was circled, was randomly but evenly assigned within each set. This resulted in a 2 (connection type: principled vs statistical) × 2 (match type: shape vs property) within-subject design, as illustrated in Table 6.2. Note that every block was randomly assigned but showed both of the sets. As explained above, within each set, all four animals were shown with either the shape or the property match.

Table 6.2: Reverse paradigm of *kevtas* studies: experimental design

Block	Animal Set	Connection Type
1	A	principled
	B	statistical
2	A	statistical
	B	principled
3	B	principled
	A	statistical
4	B	statistical
	A	principled

Hypothesising that IS generic subjects are preferred for PCs might lead participants to choose this wording for trials where the multiple-choice sentences express principled connections and the sample image that is circled is the *property* match. By contrast, if the *shape* match is circled, I hypothesise that the control sentence with the specific subject form is chosen instead.

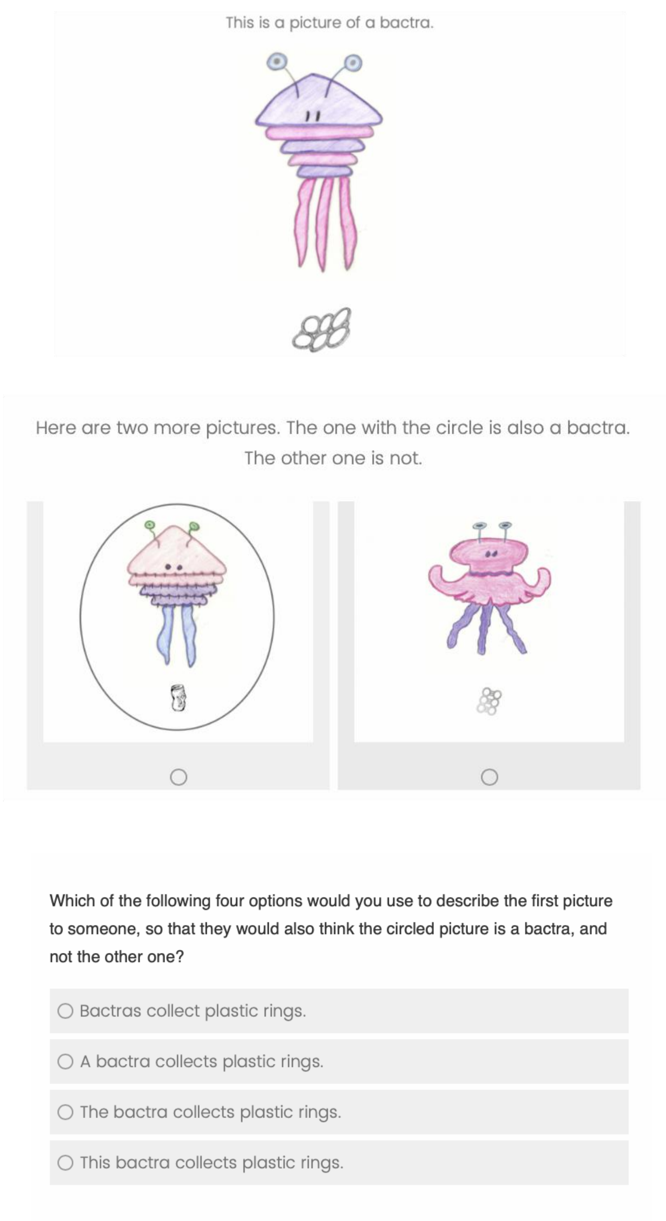


Figure 6.2: Sample prompt of reverse paradigm trial:
The target image of *bactras* is at the top, the *shape* option without plastic rings (left) is circled, the *property* match with plastic rings (right) is not. The response options at the bottom express statistical connections.

The possibility that DS subjects may directly refer to kinds leads to the hypothesis that for the trials with PCs, participants will choose this subject type for *property* matches. If the *shape* match is circled, I hypothesise that they will choose the specific subject *this*.

The versatile nature of BP generic subjects, which can express both principled and statistical connections, could be used for any of the combinations laid out for the IS and DS subjects. However, I expect that the explicit alternative of these other subjects will lead participants to choose those instead of the BP, which they might normally rely on as a default option. Thus, the BP should be chosen at a higher rate for SCs than IS and DS subjects.

6.4.1.3 Procedure

Participants were recruited via `https://www.prolific.co` (formerly https://-www.prolific.ac), which provided a brief overview of the study. They were then redirected to the study, which was hosted on `https://www.qualtrics.com`. After providing informed consent, they provided their Prolific ID and read through the following, more detailed instructions:

> Welcome!
>
> In this task, you will be seeing pictures of some new animals. They are different from the animals we know, so you will learn the name for each of them. For each animal, you will be given its name and see a picture of it along with a smaller image of something in its environment.
>
> Then you will see two more pictures. One has a circle around it and is the same kind of animal as the animal in the first picture. The other one has no circle and is not the same kind of animal.
>
> You will be asked to choose a sentence that best describes the first image so that someone else would know which of the other two animals was the same kind as the first one.
>
> Please answer to the best of your ability.
>
> When you're ready, please turn off any distractions, such as your cell phone, for the duration of the experiment! Please avoid using the mobile version of this site.

During the experiment, participants chose one of the four multiple-choice answers that they thought best described the scenario that was presented to them. They differed in the form of the subject, which was either a bare plural, an indefinite singular, a definite singular, or a specific demonstrative (*this*), as illustrated above in Figure 6.2.

After responding to the survey, they were asked to answer 6 demographics and socio-linguistic questions (see Appendix C.2) about their background and fill in an optional essentialism questionnaire of ten questions (see Appendix C.3). These were the same as those for the *kevtas*-studies. They were also given the opportunity to provide any feedback in a comment box before completing the study.

6.4.1.4 Analysis and Coding

The data of 161 participants were analysed. Their responses were converted from absolute counts of choosing a specific wording condition to probabilities of choosing a specific wording response in any condition relative to the other three wording options.

A 2 (match type: *shape* vs *property*, within participants) × 2 (connection type: PC vs SC, within participants) ANOVA was conducted to establish whether the differences in the probabilities of choosing one specific response out of four possible multiple-choice responses (with a BP, IS, DS or 'this' subject) based on the condition in which they were presented were significant. This analysis revealed a main effect of match type ($F(1,1244) = 19.458$, $p < .0001$), but no main effect of connection type ($F(1,1244) = .265$, $p = .5427$). There was no interaction of match type and property type ($F(3,1244) = .124$, $p = .7509$). These effects are illustrated in Figure 6.3. Due to lack of interactions in the omnibus model, no planned comparisons were carried out.

Instead, I conducted nested post-hoc paired comparisons, comparing *this* vs generic subjects, PL vs SG generic subjects, and then DEF vs INDEF singular generic subjects. These revealed that all generic conditions differed from the baseline *this* ($p < .0001$). Looking at the subset of generic subjects, the BP differed from both IS and DS subjects (both $ps < .0001$). Looking at just the singular generics, IS and DS subjects did not differ from another ($p = .848$).

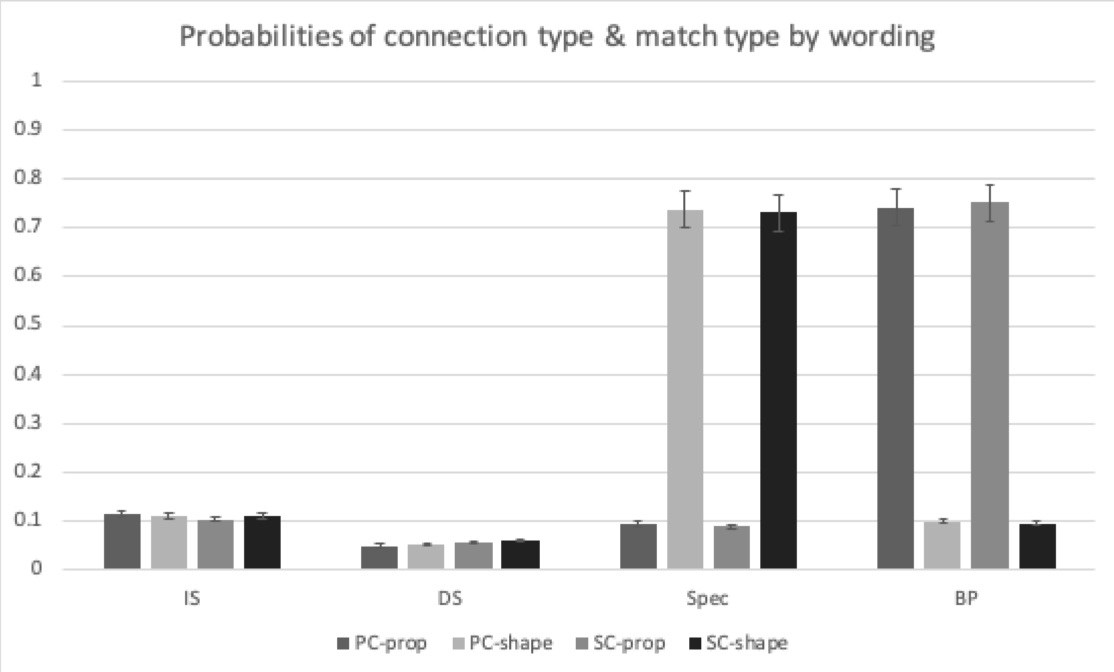


Figure 6.3: Probabilities of connection type & match type by wording for Experiment 14 (production). Error bars = 95% Confidence Interval.

I also conducted an analysis of overall response times. Unfortunately, individual reaction times per item were not available such that this is a preliminary analysis. A oneway ANOVA was conducted to establish whether the differences in response times (in seconds, per participant), based on the wording chosen from the 4 multiple-choice options (subject form: BP vs IS vs DS vs 'this', within participants), were significant. This analysis indicated a significant effect of subject form on response times ($F(3,1244) = 6.4403$, $p = .0003$).

Planned comparisons (Tukey HSD) show that, compared to the DS, participants were significantly faster to respond with a BP subject ($p = .0003$) or the specific control condition 'this' ($p = .0007$). There were no other significant differences between different subject forms (all p values between .1099 - .9689). All response times are summarised in Table 6.3.

Further analyses of reaction times showed an interaction of responses and condition ($p = .0008$). Due to the large number of levels within pairwise comparison, I will only report a subset of the conditions that differed in response times. The biggest differences were seen between DS responses in PC-*shape* conditions and

Table 6.3: Experiment 14 (production): Mean response times by wording condition

Level	Mean	Std Error	Lower 95%	Upper 95%
BP	408	6.664	394.53	420.68
IS	436	12.879	410.93	461.46
DS	486	18.214	450.71	522.18
'this'	412	6.715	398.72	425.07

'this' responses in SC-*shape* and PC-*shape* conditions (p = .0034 and p = .0315, respectively) as well as BP responses in PC-*property* and SC-*property* conditions (p = .0044 and p = .0126, respectively). I will discuss potential implications of these differences in more detail in Section 6.4.1.5 below.

6.4.1.5 Results and discussion

The most noticeable result from the reverse *kevtas* studies is that most participants opted for BP subjects, and rarely chose either IS or DS subjects. This is in line with observations in corpora of child-directed speech as well as the ubiquitous nature of BP generic subjects in naturalistic speech. These results seem to, in the first instance, confirm suspicions that BP subjects are the preferred generic subject. Yet, the motivation for this study was to tease apart the conditions under which an IS or a DS subject is preferable, both one over the other and to a BP, an observation which is supported by linguistic intuitions in the literature. Still, no differences were found between the two singular subjects.

I have also mentioned differences in response times in the analysis above. While the focus of this study was not to compare reaction times, and did not collect per-item reaction times, looking at the differences can still point towards a deeper understanding within the wider discussion of default generalisations. Future studies should explore the individual reaction times per wording condition to conduct more comprehensive analyses of the interaction of subject form and connection type of the predicate. Recall that under *Generics-as-Default*, we should expect that generic sentences as the default, either as a whole or a subset of them, should be processed more quickly than other types of sentences. This is because non-default expressions would be expected to require more cognitive effort than cognitive defaults. For this experiment, the overall fastest reaction times were seen for the control condition *this* and BP generics. These were then followed first by IS subjects, and then DS subjects, which were the slowest. The interaction between responses and condition were driven by the slower responses of DS options in the PC-*shape* condition.

Bearing this in mind, there can be multiple explanations as to why response times differed. Since these studies were not designed to test speed, or targeting a speed-accuracy trade off, this post-hoc reasoning must be taken with a grain of salt. In terms of choosing generic sentences based on the connection and match type, the data showed that the singular generics are the same, but that they differ from the plural generic subject. Returning to the initial hypothesis, I argued that singular generics should be preferred for PCs, and that the DS should be chosen at a lower magnitude than the IS. While these differences were not reflected in the probabilities of choosing a specific sentence, the differences in response times show that the IS and DS are distinct from the plural generic, and that when the DS is chosen, it takes participants longer to do so. If we take faster reaction times to reflect, or correlate with, a cognitive default, then the first conclusion would be that if generics are a default, this cannot apply to all morphosyntactic forms of generics. Based on this argument, the BP generic should be the default generalisation, and the IS and DS are either not a default, or are hierarchically below the preferred BP default generalisation.

If, generally, faster reaction times indeed reflect a default, then the control condition *this* could equally be treated as such. This would allow us to assume that both BP generics (and specifically not IS and DS generics) and the specific demonstrative *this*, are cognitive defaults.[10] One explanation for this could lie in the two types of interpretations that these render, i.e. the difference between interpreting the response sentences generically and episodically. However, this study did not specifically target the interpretational differences between BP generics and episodic

[10]Seeing that I am distinguishing between different types of generic sentences, I equally do not want to group together all demonstrative determiners. This mirrors earlier discussions about using quantified statements as baselines against which to compare generic sentences. It seems that *all*, *all the* and *each* do not lead to the same inferences, thus I will not assume that *this* would lead to the same assumptions as its plural or distal variants, e.g. *these, that, those*.

'this' statements and it is therefore difficult to judge whether there are two preferred, or 'default' wording options, one for generic and one for episodic statements.

With this caveat, let us briefly expand on the discussion of reaction times to raise the possibility of an alternative that may be explored further in future research. One could also argue that none of the responses that scored faster across the board are the cognitive default. This would be in line with the argument that acquiring the varying truth-conditions for exception-tolerant generics is difficult, and possibly more difficult than learning about specific instantiations in spatio-temporally defined contexts (e.g. due to a lack of overt marking, see Reiter & Frank, 2010 for a computational linguistics/machine learning approach or Teichman, 2015 for a philosophical approach on generic quantification). However, two issues with this theory come to mind. First, how do we account for the fact that we have seen in Section 4.4 that generics, whether singular or plural, are understood and interpreted correctly early (as also discussed in many other psychology studies for BP generics only). In those same experiments, children start out with high *property* matches, seen as an indicator for generic generalisation, for *this*? Those scores gradually decrease and become more in line with the results seen for adults, in which *this* acts as a more appropriate baseline and is preferred for *shape* matches instead. Second, how do we explain the developmental differences that we have observed in these studies between plural and singular generics, the latter of which seems to be acquired gradually and later? On the one hand, the gradual and later acquisition of singular generics provides a nice parallelism with the slower reaction times compared to plural generics. On the other hand, these findings cannot be taken as conclusive in terms of which, if any, of these forms are more default than others.

This leaves us with two options to consider for the follow-up studies in Section 6.4.2 below. Either, reaction times are not a good measure to address the question of defaults in the human conceptual system. If the patterns for the dependent variables and the reaction times in Experiment 14 (production) are the same, then a more thorough investigation of this correlation is warranted. Alternatively, we should keep in mind that *this* and BP subjects have distinct underlying mechanisms, which have simply been conflated by the methodology in these studies. Returning to teasing apart SG and PL generics, Experiment 15 (production) is specifically designed to move away from relying on BP subjects. If we consider how to better understand the role of these two singular subjects, the most straightforward option would be to remove BP subjects from the answer options, and thereby forcing participants to choose between singular subjects if they wanted to use a generic sentence. This option is presented in Experiment 15 (production) below.

6.4.2 Experiment 15 (production): Reverse paradigm without BPs

Experiment 14 (production) showed that participants in a production-like study preferred to use BP subjects for *property* matches, regardless of connection type, and *this* subjects for *shape* matches, also regardless of connection type. Neither IS nor DS subjects were chosen frequently. Experiment 15 (production) follows up the hypothesis of distinct semantics leading to distinct behavioural results by removing the option of a BP subject, thereby forcing participants to decide between IS or DS subjects, should they want to choose a generic statement.

These experiments include a new response option, *None of these*. This allows participants to choose a response that is not only neither a singular generic subject nor a specific subject that allows for individual, non-generic reference, but is intended to give participants the opportunity to respond without choosing any of the available options, in case they strongly prefer a response that was previously available via the BP, seeing that the stimuli are the same.

6.4.2.1 Items

The response options with BP subjects were replaced with the sentence *None of these*. All other materials were the same.

6.4.2.2 Participants

160 monolingual, native speakers of American English were recruited via Prolific under the same conditions. None of them had participated in Experiment 14 (production). They were randomly assigned one of 4 conditions (n $\approx$ 40 per condition).

6.4.2.3 Procedure

The procedure was the same as in Experiment 14 (production).

6.4.2.4 Analysis and Coding

The analyses were carried out as in Experiment 14 (production).

As for the previous experiment, the data were converted from absolute counts to the probability of choosing a specific wording response in any condition relative to the other three wording options.

A 2 (match type: *shape* vs *property*, within participants) $\times$ 2 (connection type: PC vs SC, within participants) ANOVA was conducted to establish whether the differences in the probabilities of choosing one specific response out of four possible multiple-choice responses (with an IS, DS, 'this' or *None of these* subject) based on the condition in which they were presented were significant. This analysis indicated a main effect of match type ($F(1,1074) = 147.609$, $p < .0001$) and of connection type ($F(1,1074) = 3.45$, $p = .0003$). There was also an interaction of match type and property type ($F(3,1074) = 2.419$, $p = .0038$). The probabilities of choosing a specific wording condition are illustrated in Figure 6.4.

Planned comparisons showed that for *None of these*, SCs were preferred over PCs ($p = .005$). For the IS, PCs were preferred over SCs ($p < .0001$). Further, for the DS, *property* was preferred over *shape*, regardless of the connection type expressed in the answer options. For *None of these shape* was preferred over *property*, which shows the opposite pattern as the DS.

Nested post-hoc subset analyses were also conducted for this experiment, comparing the means of the responses. The initial division was between *None of these* and the remaining three response options, based on the distinction that the

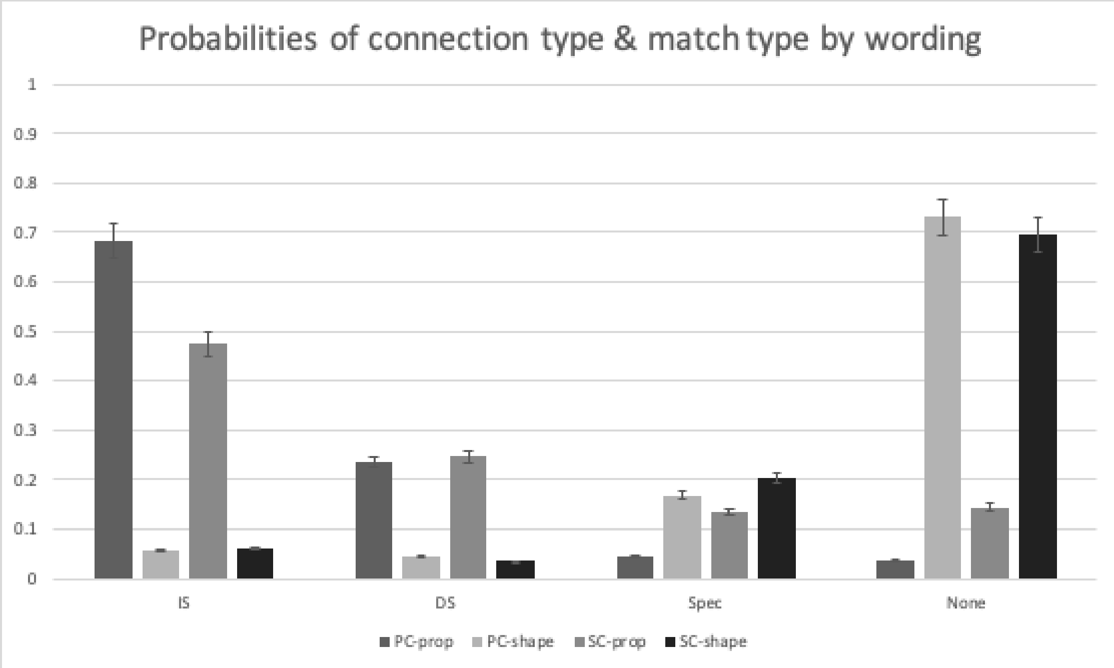


Figure 6.4: Probabilities of connection type & match type by wording for Experiment 15 (production). Error bars = 95% Confidence Interval.

remaining three were all full sentences that participants could choose to describe the triplet of images to someone else. The next division was between non-generic and generic responses, i.e. *this* and the IS and DS responses. The final division was between the IS and DS subjects. An ANOVA shows that comparing *None of these* with the three full-sentence responses, these response options differ from each other ($p < .0001$). Looking only at the labelled options, *this* also differs from the two singular generic subjects ($p < .0001$). Finally, IS and DS subjects do not differ significantly ($p = .077$), although note that this is marginally significant and that they differ to a much larger extent than in Experiment 14 (production) (where IS vs DS rendered $p = .848$).

I also conducted an analysis of overall response times. Note that as for the previous experiment, reaction times for individual items were not collected. Thus, the analysis has the same exploratory and preliminary character. A oneway ANOVA was conducted to establish whether the differences in response times (in seconds), based on the wording chosen from the 4 multiple-choice options (subject form: BP vs IS vs DS vs 'this', within participants), were significant. This analysis indicated a significant effect of subject form on response times ($F(3, 1074) = 3.1644$, $p = .0238$).

Planned comparisons (Tukey HSD) show that, compared to *None of these*, participants were significantly faster to respond with *this* ($p = .0198$). There were no other significant differences between subject forms (all p values between .2644 - .9983). All response times are summarised in Table 6.4.

Table 6.4: Experiment 15 (production): Mean reaction times by wording condition

Level	Mean	Std Error	Lower 95%	Upper 95%
IS	411	15.454	380.32	440.97
DS	415	23.071	370.08	460.62
Spec	364	25.628	313.71	414.29
None	448	13.306	421.65	473.87

One conclusion we might draw on this ranking is that participants considered all other options before resorting to *None of these*, thus in fact using it as a last resort when no other labelled option was deemed appropriate.

6.4.2.5 Results and discussion

One of the most striking findings from Experiment 15 (production) is the strong preference for *None of these* for the *shape* options, regardless of whether the response options described a PC or SC. This is the pattern previously found for the specific baseline *this*, which was the least preferred option in these studies. However, a direct comparison of these two types of baselines is difficult as Experiment 14 (production) did not include a *None of these* option alongside the *this* baseline. These results partially support the hypothesis that linguistic form matters in the conceptualisation of novel kinds and their category-property links. Once only IS and DS subjects are available, they display the pattern of the previously preferred BP (in Experiment 14 (production)). Interestingly, the newly introduced option *None of these* displayed the pattern that was found for *this* in Experiment 14 (production). Thus, while *None of these*, in a sense, took on the pattern for *this*, in this version of the studies *this* was rarely chosen. In the cases where it was chosen, it was done so mostly for the PC-shape options, and least for PC-properties.

This raises a new question of whether *this* is an appropriate baseline for these tasks. Parallel to preferring BPs, if they are available, and else resorting to IS and DS subjects, participants might prefer not to use any of the options, but will resort to *this* if no alternative is available.

In addition to that, Experiment 15 (production) provides further evidence that the results of production-like studies can differ from comprehension ones. This becomes particularly clear when comparing the distinct patterns of singular and plural generic subject to the original match-to-sample studies in Chapter 4. While in those earlier studies, participants often chose *property* matches regardless of whether they were presented with a BP, an IS or a DS subject, participants in this reversed paradigm were most likely to choose IS subjects for *property* matches (without the option of choosing a BP). By contrast, they mostly chose *None of these* for *shape* matches. The third most probable subject form that was chosen for *property* matches was the DS, while *this* was rarely chosen at all.

Thus, the first general point from this version of the reverse match-to-sample task is that *this* might not be the best baseline for match-to-sample studies. At least when it comes to matching members of animal kinds, it is conceivable that *shape* is already in itself a principled property. This means that these studies conflate one type of PC (*shape*) with another, explicitly mentioned PC (e.g. *have stripes*). Another comparison that has not been addressed explicitly regards the expectations of the relative strengths of PCs vs SCs. This would be relevant when considering the differences between the PC-shape and SC-property conditions, in which one might hypothesise that the explicitly mentioned *property*, albeit statistically connected, might be relatively less salient than the *shape* option which co-occurs with a PC property.

There is a question as to how transferrable the observation of a potentially mis-chosen baseline for this task could be for the original match-to-sample tasks. I argue that the concern regarding *shape* as a baseline is valid for both the experiments in this chapter, as well as for the Experiments 7-12 (world knowledge) in Chapter 4. The reason is that for both tasks, generalisations were made at least partially based on visual information. In turn, this provided participants with knowledge about the category type, which was an animal kind. Based on the discussion of overhypotheses in Chapter 5, I assume that participants would assume certain characteristics to hold true of the animals, since they are true of other similar animals. These similarities are predicted of novel animals via connections that are assumed to be inherited via superordinate animal kinds. This holds regardless of whether the role of morphosyntactic cues can be teased apart more in a comprehension match-to-sample task or a production-like reverse match-to-sample task.

The next section will compare the results of the two reverse match-to-sample studies before then turning to a general discussion of the findings of this chapter.

6.4.3 Interim conclusions

Comparing the results of Experiment 14 (production) and Experiment 15 (production), it can be argued that when all generic subjects are available, the BP is indeed preferred in production. This is because when provided with the option of the plural generic, singular generics were rarely chosen. However, without the BP subjects to choose, IS and DS subjects show high scores for *property* matches, which was previously observed for BP subjects. Moreover, the IS shows an interaction, preferring PCs over SCs, in line with the hypotheses in the semantics literature.

Another argument one might pursue is that the studies simply provide evidence that the BP is preferred over any SINGULAR generic. However, we have seen before that both IS and DS provide easy access to generic interpretations in comprehension studies. The observed difference between PLURAL and SINGULAR generics could be attributed to a difference in the lower frequency of producing SINGULAR generics. This would be supported by the corpus analysis and picture book analyses in Chapter 2, showing that while IS and DS generic subjects are not absent, they are much less likely to be used to introduce novel or existing kinds. This is in line with independent reasoning in which an out-of-the blue DEFINITE DP is unusual to introduce a novel individual or kind. DS subjects are proportionally far behind BP subjects, PLURAL pronouns referring back to BPs, and quantified statements, although the exact ratios are dependent on the type of category that the picture book introduces (e.g. animal kinds vs artefacts).

Even considering these more basic refinements of the methodology that was tested here for the new reverse match-to-sample tasks, these studies provide an insight into an as-of-yet under-explored way of exploring the differences between generic comprehension and production. Providing participants the opportunity to compare multiple generic subjects asks them to tap into their meta-linguistic awareness in a way that a between-participant design and a more traditional comprehension match-to-sample study does not. Going forward with these studies, there are some changes that I would implement to gain a better understanding of a methodology that asks participants about their preferred generic phrases:

1. Seeing that in many cases, English generic statements allow for multiple generic subjects, giving participants the opportunity to either choose more than one of the responses, or even to rank them in order, would allow for a more accurate depiction of participants' judgements. Especially since Experiment 14 (production) showed no preferences for IS or DS subjects, giving participants the option to choose more than one response could show whether they were simply not an option, or alternatively just not the most preferred option.

2. To address the puzzle of participants choosing *None of these* over *this* in Experiment 15 (production), a follow-up experiment could provide participants with both *this* and *None of these* as non-generic response options in a version where the BP is still an option as well, thus ultimately giving participants five instead of four options from which to choose.

3. As there were only minor differences between *shape* and *property* options for most generic subjects (except for the interaction within IS subjects in Experiment 15 (production)), another follow-up experiment to establish a baseline account for this reverse paradigm could involve a simpler version of this task. Participants could be given just the PC and SC phrases, but no

shape matches, thus only focusing on the differences between statistical and principled connections in generic generalisations.

4. Taking into consideration Lazaridou-Chatzigoga et al. (2019)'s findings of context-sensitivity for generic statements, the results could be affected by providing more context for a setting in which participants are asked to make these judgements. One context could be imagining talking to a child and instructing them, similar to popular picture book studies.[11] Taking the studies in Chapter 4 (with child participants and audio-stimuli) as an indicator, this could lead to higher probabilities of participants choosing DS generics, as they are often perceived as instructional, and more appropriate in an educational setting.

5. Finally, having used the high essentialism assumed of animal categories as an advantage to streamline the categories tested in these studies, as well as to ensure that certain properties are likely perceived as tightly connected to the kind, another possible comparison could include different categories, such as social categories. Working under the assumption that DS subjects, as well as BPs and even DPl subjects are commonly employed to refer to social groups, this might give an entry into exploring other distributions. I would, however, proceed with caution as the specific categories and properties can also be associated with negative propaganda and I therefore expect that participants might be more reluctant in ascribing properties with generic language in such a setting.[12] More neutral cases of social category ascriptions, such as *Dutchmen are good sailors* have been raised and discussed as recently as in Nickel (2012).

It is important to keep in mind that the *shape* and *property* conditions express multiple biases. This has previously been used as an advantage (initially in Hollander,

[11] I thank Dimitra Lazaridou-Chatzigoga for pointing this out to me during discussions at the Cambridge Generics Workshop held at Pembroke College in January 2020.

[12] I thank Rae Langton for making me aware of such cases, particularly thinking of historical cases in which these generalisations have been used politically, during the Cambridge Generics Workshop.

2007). To test these assumptions, one potential alteration of this design could see crossing the overall design. To better explain this alternative design, recall the novel animal *bant*. The SC property was *taps bottles* and the PC property was *has three legs*. In the previous studies, the match with three legs was dissimilar in shape to the target and was also accompanied by a little image of a bottle since the images were the same, regardless of whether the participant was in the PC or SC condition when seeing the triplet of images for the bant. Now imagine that instead of the PC property *has three legs* being accompanied by the image of the bottle, it was instead accompanied by the non-match of the SC condition, e.g. a wooden bowl. I hypothesise that in the PC condition, the participant's attention would not explicitly be drawn to the little image. Yet, this set-up would provide a stronger competition with the *shape* match, as the non-property option is both more similar in overall shape and also has the same type of artefact associated with it (although the artefact would not be explicitly mentioned in this condition). The problem with providing this stronger baseline test, to overcome the *shape* match, becomes apparent when this set-up is applied to the SC condition. In this condition, the relevance of the artefact is the crucial variable that is manipulated. This means that providing the participant with the predicate *taps bottles* draws attention to the accompanying image. Here, the design that I suggested for the PC condition would not work in the same set-up, as both the predicated property and the *shape*, which functions as a baseline, would be the same image. Participants would therefore never be drawn to choose the option that looks both less similar to the target animal and also does not have the predicated property. This could be a stronger test for the PC condition, to counter the inherent *shape*-as-a-principled-property bias, but would then mean that the SC triplets are different. Any differences between PC and SC conditions could then either be attributed to differences between the triplets themselves or to the manipulated *shape* vs *property* variables.

Overall, the fact that singular generic subjects display distinct behaviour in this follow up is noteworthy for a more general reason: The same visual stimuli have been used as in the studies in Chapter 4, where these more fine-grained differences were not observed. Consequently, we might wish to revise the hypothesis that attention to linguistic differences is limited when visual information is present. Instead, this overriding effect of visual stimuli might be limited to perception-like studies or, more generally, be limited by the experimental methodology. Reversing the match-to-sample methodology has shown to be a way to target production of generic sentences, forcing a decision between (singular) generic sentences within one participant. It has also shown that, if given the opportunity, the BP remains the preferred option as it allows for the widest range of semantic representations and is ubiquitous in its use.

6.5 General discussion

This chapter started by taking a step back and looking at methodological differences and general categorisation strategies that have been proposed and adapted independent of this research. Based on an evaluation of their different assumptions, a more targeted review of the different experimental results concluded that the different methodologies that I have applied led to distinct results, which either enhanced or conflated the morphosyntactic differences under review. Consequently, I used this new-found understanding to my advantage by implementing a novel design, created to tease apart the differences between various generic subjects.

The studies in this chapter are methodologically unusual in that they provide participants with category information and ask them to give instructions such that others would come to the same categorisation decisions. While this has led to novel insight in terms of production decisions of generic subjects, it also seems to present participants with an unusual task by calling on their metalinguistic awareness. To counter this specific argument, recall that Plunkett (2015), Plunkett & Sundell (2019), and Sterken (2020) have argued that metalinguistic negotiations are pervasive in everyday speech. A prerequisite to use metalinguistic negotiations is to have some form of metalinguistic awareness, whether this be implicit or explicit. This goes back to ideas proposed by ter Meulen (1986) and Cavedon & Glasbey (1994) within the framework of Situation Theory, in which generics are evaluated by the constraints they impose on situations (as opposed to constraints on individuals as in normalcy-based semantics, e.g. Chierchia, 1995; Declerk, 1991, although see Leslie & Lerner, 2016, Section 2.2 "Relevant Quantification and Situation Semantics" for a more comprehensive overview).

Suppose that because this is a rare task type, participants did find it more difficult and perceived the task to be artificial. In fact, based on the comments that were provided by some participants at the end of the study, many of them indeed found the task to be difficult and artificial. However, this was a common comment for the other experiments in this thesis as well, some of which were replications of well-established experimental paradigms. This is not to say that these are therefore unproblematic. To the contrary, I hope that the discussion sections for the experiments have proved insightful in terms of how the design could be more reflective of actual language use and how methodology should be considered as a significant factor influencing experimental results. Moreover, these comments were overall balanced with comments from other participants that said they found the study fun, easy or interesting. Of course, some comments that come up multiple times are taken seriously for follow-up experiments. But it should also be considered that even highly unusual studies, such as the *pseudoword* studies in Chapter 3, after which participants commented that they were very confused and did not know whether they understood the task correctly, provide results that were consistently found in replications and follow-ups. Thus, to draw on linguistic knowledge, participants need not find the task intuitive or natural.

The non-intuitive design could be interpreted to particularly disadvantage the *shape* condition. Participants are asked to decide between four sentences, all with the same predicate, e.g. the principled property *are striped*. They could then be in a condition where the predicated property that is repeated in all four multiple-choice options is the property that is in the non-circled option, e.g. the picture of the novel animal that is not striped. In other words, the only characteristic that is predicated in the response possibilities is the one that is depicted in the choice which the participant is told is **not** an instance of the novel kind. This lack of overlap in both the visual and linguistic information provided could explain the high proportion of the *None of these* responses in Experiment 15 (production), which only occurred for the *shape* condition.[13]

Additionally, Chapter 4 looked in more detail at the original match-to-sample studies by Hollander et al. (2009). One of their main conclusions was that "[s]urprisingly, **children also selected the predicated-property choice significantly more often in the non-generic condition than the control condition**. (...) Thus, despite the strong pull toward the shape-match when only perceptual information was available (i.e., the control condition), **children switched to selecting the property-choice match more often in both the generic and non-generic conditions**" (Hollander et al., 2009, p. 497, emphasis added). Further research should initially revisit some of the assumptions made in Hollander's design. For instance, one reason for the high *property* matches for *this* could be that *Generics-as-Default* (Leslie & Gelman, 2012) is at play. This offers a straightforward explanation of the developmental patterns in the child studies where the baseline gets high scores for control sentences. However, note that adults do not normally display these patterns. In Experiment 14 (production) and Experiment 15 (production), *this* did not just show high *property* matches, but rather distinct patterns based on the other options provided as multiple-choice answers.

Alternatively, high *property* matches for *this* could be due to aforementioned labelling effects, i.e. the idea that by providing a label for an object, a category is established. If, in addition to this, properties are predicated of the newly labelled object, these could lead to categorisation based on those properties being present and override perceptual similarity (as discussed in Section 5.1, with reference to Plunkett et al., 2008). This analysis would rely on the semantics of nouns which have been argued to evoke kind reference in itself (e.g. Dayal, 2004, who argues that nouns are ambiguous between kinds and objects; Borik & Espinal, 2012, 2015, who argue that nouns are properties of kinds). By contrast, one would not assume that a category is stipulated based on visual input of one individual alone, i.e. without being given a label. This explains the results of previous studies, such as Hollander et al. (2009).

[13] I thank an anonymous reviewer for pointing out these further pragmatic considerations.

These studies do not lend themselves to testing whether a condition in which no labels are given would lead to lower *property* matches for non-generic subjects, since the set-up of the study relies on providing multiple-choice options as responses.

In summary, this chapter has highlighted the importance of better understanding the intricate interactions between methodology, stimuli type, and linguistic variation. At present, their consideration in experimental design has not been sufficient to make claims about the organisation of generic knowledge in our conceptual system. A better understanding of how these affect behavioural responses to studies targeting genericity will then allow us to make clearer judgements about the nature of *Generics-as-Default* and related theories, which are proposed to be the foundation of the human conceptual system. Note that in order to focus more specifically on the question of *Generics-as-Default*, experiments should be set up with the specific goal of collecting reactions times. This was not the focus of the studies presented in this thesis, yet it has become clear through the discussion of methodological issues in this chapter that this is an important consideration for future experimental design.

This chapter has moved away from relying on comprehension style experiments. It has provided participants with category knowledge, as some other studies in this thesis have, while reducing the amount of contextual information to match that provided in the original *kevtas* studies. I have argued that with a careful empirical approach that includes different kinds, with a variety of linguistic and visual cues, and a varied amount of exposure, extensions of the existing experiments can indeed shed light on the role of language in category acquisition. The results of Experiment 15 (production) showed patterns for the two singular generic subjects that were similar to those of the previously tested BP subject in Experiment 14 (production), thus replicating patterns within the linguistic mode of production. Most notably, the main effect of connection type, i.e. principled connections are preferred over statistical ones, was replicated in this linguistic mode.

The findings regarding the amount of exposure (from Chapter 3), the different types of baseline quantifiers (from Chapters 4 and 5), and task type (ranging from ratings on a Likert-scale to responding to questions about a picture book) confirm this repeated preference of connection type, regardless of the morphosyntactic form of the subject. Regarding the amount of exposure specifically, Experiment 1 (DEF-NUM), with medium amount of training, and Experiment 4 (DEF-NUM), with extended training, showed the hypothesised interactions, but Experiment 3 (DEF-NUM), with no training phase, showed an interaction in the opposite direction than hypothesises. Relatedly, regarding the type of information available specifically, Experiment 9 (world knowledge) showed the hypothesised interaction after excluding a prompt for participants that included a BP generic subject, as did both of the subsequent studies with adult participants (Experiment 10 (world knowledge) at a marginally significant level, Experiment 11 (world knowledge)), but not Experiment 12 (world knowledge) with child participants.

This demonstrates that the results are overall ambiguous in terms of their support for the predicted distinct behavioural responses across the studies presented in this thesis. With this in mind, understanding the precise contribution of each of the seven variables in Section 6.1 will be crucial in future experimental design and to test the hypothesised interaction of subject form and connection type in a more targeted manner.

*Guy who invented **the piano**: 200 hundred [sic] years from now **it** may need tuning but **it** will be sturdy. So sturdy.*
*His friend, who invented **piano benches**: the legs are designed for maximum wobble[.]*

— @LlamaInaTux *Twitter, Feb 15, 2019*

7 Conclusions

Contents

7.1 Linking linguistics and psychology

This thesis has investigated views on genericity held in linguistics, complemented by approaches from psychology, to arrive at a more unified theory of the role of generics in concept acquisition and processing. The inclusion of the psychology literature offers a unique perspective to psycholinguistic approaches, as it has argued that generics are the cognitive default, seeing that they are processed more quickly. Linguistic arguments, on the other hand, have treated generic subjects as a heterogeneous group with distinct meanings. In line with this view, I have started by motivating four distinct layers of the DP (NP, ClP, #P, and DP) to illustrate the differences in NUMBER and DEFINITENESS that account for the interpretational differences of the three generic subjects for count nouns in English. These differences motivated the experimental methodology for the subsequent experiments, thereby combining methodologies and findings from linguistics and psychology.

Following the assumption that nouns denote properties of kinds, this thesis made various inferences about the role of the noun in generic subjects. Firstly, if the root of a generic subject makes reference to kinds, this may provide an explanation for labelling effects, which have long been observed in developmental psychology. To illustrate, when presenting a child with a novel member of a kind, e.g. a novel animal or artefact, they perform better at memory tasks, such as recall, and categorisation tasks if this novel item is introduced with a label. If we assume that a subject DP evokes kinds, these labelling effects are a consequence of referring to a (new) kind, and thereby establishing the novel kind's category structure. Secondly, this explains the special status of kind-referring subjects within the various types of generics. There are specific kind-level predicates which are best predicated of kinds expressed by the DS (or, as reinterpreted, the DN). The kind-referring DS is also the subject form with the most limited distribution.

Moreover, the differences in the layers of the DP, and in particular the values of the #P projection, help to understand the distinct meanings and uses of IS and BP subjects. By contrast, psychologists have rarely included IS subjects in their experimental design (although see Rhodes et al., 2012), arguing that they are similar to the BP but generally express more essential properties. The DS has been included in even fewer studies addressing genericity.

To recapitulate, semantically, IS subjects refer to arbitrary individuals that are expected to possess the predicated property. As such, IS subjects are generally thought of as expressing principally connected properties. However, the methodology of these studies did not find an interaction of subject type and preference for connection type. Rather, a general preferences for PCs over SCs was found, but this was not always significant and neither was this limited to IS subjects.

BP subjects represent groups and their properties are shared via join semi-lattice structures. As such, BP subjects are generally thought of as expressing either principally or merely statistically connected properties. In the studies, they supported both PCs and SCs. Crucially, though, BPs supported PCs more than SCs, similar to IS subjects, and therefore did not express these connection types to an equal extent.

For DS subjects, the experiments provided support that they are typically interpreted generically, as opposed to episodically. While they displayed similar patterns to the other generic subject overall, the results were lower in magnitude than IS subjects. This was predicted and is possibly due to their low prevalence. Crucially, they supported both PCs and SCs, to a high extent, e.g. more than the baseline 'this' for the experiments that used this as a baseline. Moreover, similar to BP and IS subjects, they displayed a general preference for PCs over SCs.

These observations led to the initial assumptions of the experimental work in this thesis: Differences in the morphosyntax of generic subjects should be indicative of differences in their underlying semantic representations. In turn, these should lead to distinct behavioural responses, based on the distinct contexts in which they are used. In other words, my experiments were motivated by the initial assumption that these generic subjects fulfill distinct roles for the human conceptual system. Alongside the expectation that distinct generic subjects lead to distinct behavioural responses, my motivation for running the experiments I presented in this thesis were two-fold:

First, linguistics and psychology have proposed two different types of theories that should account for the processing of generic statements, *Generics-as-Default* and GEN. This thesis questions whether *Generics-as-Default* can account for all occurrences of generic statements or whether the morphosyntactic make up of the subject modulates our expectations of categories and their properties. The experiments in this thesis aimed to revise the evidence that has been brought forward for these two theories to ultimately contribute to the proposition of one unified theory and were specifically motivated by the observation that English uses different types of generic subjects. The hypothesis that I proposed initially concerned whether, if manipulated carefully, these distinct generics evoke differences in behavioural responses.

The results of the studies in this study have provided intransparent evidence for this hypothesis. Some experiments have shown distinct treatment for each generic subject, while others did not. Especially in the comprehension studies in Chapters 3, 4, and 5, there were no reliable and consistent differences in participants' responses based on the wording of the generic subject in line with the hypothesised interactions between subject form and connection type.

Notably, all experiments showed a main effect of connection type, favouring PCs over SCs, regardless of the morphosyntax of the subject. This is particularly notable for the BP, which is generally assumed to support both connection types equally. One main takeaway message is that all generic subjects can perform very similar functions. While the experiments generally showed a lower magnitude for DS subjects, this might be due to its lower frequency. Yet, in studies that provide more training (e.g. Experiments 1-5 (DEF-NUM)), participants still treat this subject similarly, which indicates their increased willingness to accept even rare generic sentences. This assumption is supported by having observed the opposite pattern in the no training condition of Experiment 3 (DEF-NUM), which favoured the most frequent generic subject, the BP. Alternatively, this could be attributed to the different semantics of the DS alone, or to a combination of both of these variables.

Similarly, in Chapter 4, the studies consistently showed a main effect of connection type. The hypothesised interaction was only found once the initial prompt, which included a BP subject, was taken out. This leaves the question of the role of morphosyntax obfuscated, but raises the question of the role of methodology. One specific aspect that Chapter 4 addressed is the role of prosody. Seeing that morphosyntax alone does not reliably predict distinct behavioural results, the different generic subject types that I have investigated from a morphosyntactic point of view should further be considered in conjunction with a phonological perspective. This morphophonological approach could be tested by adding audio cues to the *zarpies* picture book study to see if the DS condition is interpreted generically at higher levels.

This raises another surprising result: The baseline condition in the studies in Chapter 4 only displayed adult-like behaviour gradually. Initially, *this* received similar ratings to BP generic subjects. In other words, the responses to specific subjects moved away from responses that were in line with generic subjects. One might argue that this is evidence for *Generics-as-Default*, and that specific reference needs to be acquired. However, I argue that an alternative interpretation of these results returns to the notion that nouns denote properties of kinds and therefore evoke generic reference. Future research should focus on disentangling the assumptions of *Generics-as-Default*, i.e. the assumption that the conceptual system is set up to interpret generically in the absence of clues not to do so, with the assumptions of kind reference made by nouns.

In Chapter 5, I explored the role of having access to general, high-level category knowledge. I made use of picture books with novel animals kinds and found effects of property types with regard to world knowledge, which are not driven by morphosyntax (and, in fact, treat the quantifier *most* akin to the BP and the IS). The effect of category knowledge, animal essentialism, and overhypotheses contributed significantly to the results, either in addition to or in lieu of an effect of morphosyntax. This implies that category knowledge, even if it only pertains to general, superordinate kind knowledge, can override subtle morphosyntactic cues, at least in comprehension studies. This also indicates that while morphosyntactic differences reflect semantic differences, these might not be too subtle for the conceptual system to affect behavioural responses.

Finally, the reverse match-to-sample studies in Chapter 6 demonstrate an overall preference for BP subjects, but that participants will resort to using IS and DS generics if the BP is not available. As with the previous studies, there is a main effect of preferring options that explicitly show the predicated *property* over a *shape* match, and PCs over SCs.

Therefore, the studies presented in this thesis show that adults are aware that different types of generic subjects exist, and that they employ them. This was more prominent in the comprehension studies, especially those in which wording was manipulated as a between-participants factor, but also applies to the production studies, especially those in which the BP was not available as a response option. Yet, the predicted interaction was not reliably found and it therefore remains unclear what the precise contribution of morphosyntax is to the overall interpretation of generic sentences. In terms of the discussion about *Generics-as-Default* and GEN, this thesis found no direct support for the hypothesis of a covert GEN that leads to distinct behaviour. On the contrary, the results of this thesis have shown that adults are willing to accept even rare generic subjects as generic, which is more in line with the idea of *Generics-as-Default*.

This leads to the discussion of the second motivation: The debate on *Generics-as-Default* and GEN, as well as the overall research agenda of the morphosyntactic differences of (English) generic subjects, seeks to address whether some of these representations and interpretations are readily available to young children. I asked whether both adults and children can make use of the semantic differences to make inferences about the nature of category-property links. I hypothesised that if these different generics meanings are readily to (young) children already, then they should already display (a) generic interpretations to less commonly tested generic subject types, e.g. IS and DS generics; and (b) if adults display distinct behavioural responses for each of these generic subjects, and children already have distinct representations for these different generic subjects, then they, too, should be sensitive to their meanings and display distinct behavioural responses.

The results have shown that while young children are already sensitive to generic interpretations for IS and DS subjects, like adults, they do not display distinct behavioural responses based on the specific morphosyntactic form that the generic subject has. These studies have shown that adults and children are able to interpret all three generic subject types as generic, and that even rare forms of generics are treated similarly to the ubiquitous BP. The only consistent effect that these studies have shown was a preferences for principled over statistical properties. It seems that generic interpretations are readily available, albeit their interpretation seems to be modulated to an extent by frequency of exposure. This provides further evidence for the hypothesis that generics could be the default way of conceptualising instances of kinds. This, though, does not directly refute GEN, as these interpretations might also depend on the accessibility of A-quantifiers, as raised in Chapter 2.

Taking the mixed results from all of the studies seriously, there are at least two possible reasons why the results do not support the hypothesis that is based on linguistic observations.

1. The differences in morphosyntax, which reflect different semantic representations, are differences that we are not sensitive to.
   This is a stark rejection of the underlying assumptions made in the linguistics literature. While this is possible, it would have implications that disregard the differences in linguistic judgements. As discussed in Chapter 2, this could be further explored by collecting grammaticality judgements specifically targeting generic statements.

2. The design of the experiments is not sufficient to target these differences and affect different behavioural responses.
   Based on the theoretical considerations and some tentative trends in this thesis, I propose this to be the more likely explanation. Given that the studies in this thesis are a novel way of examining linguistic subjects experimentally, it is not clear yet which would be the best experimental set up to target the hypothesised interactions. The methodology review in Chapter 6 discusses the amount of training, the right type of labelling, the use of animal kinds with high essentialism, and the task type as possible factors that contributed to the mixed results in this thesis.

The overall takeaway message is that there is no consistent evidence that morphosyntax strongly modulate category-property links. Despite the initial hypothesis that generics should not be treated as a unified phenomenon, they evoke similar behavioural responses. I have contextualised these results by highlighting other types of information that were available to participants in Chapter 6, arguing that the differences in the morphosyntax might not be crucial differences to the conceptual system, at least not when evaluating how to think about the properties that are predicated of a (novel) kind. In other words, unlike the initial hypothesis, just because these different types of generic subjects are available to the conceptual system, this does not mean that we make use of them to inform category-property links.

Future research should focus on evaluating the validity of the initial hypothesis by collecting empirical data about the acceptability of generics. The assumptions that motivated this thesis are based on introspective, individual judgements. Parallel to how this thesis argues for the inclusion of a wider range of generic subject types in psychology studies, in this case linguistics should include psychology methodology to a larger extent and collect acceptability judgements on generic sentences to further support their hypotheses about their differing distributions. However, even if these empirical data support the linguistic intuitions, it could still remain the case that while there are multiple types of generic subjects, they might not be that crucial to the conceptual system, whether without access to world knowledge in an experiment or with category knowledge. Then, further research is need to explain how linguists and psychologists have pointed out, e.g. that IS subjects are better suited to express PCs, and how this is not supported by empirical data yet.

The following sections will discuss some implications of the main findings of the studies regarding the interplay of morphosyntax and the human conceptual system. Section 7.2 addresses potential confounds in the types of cues that are available to us when making inferences about novel kinds. Section 7.3 presents one possible proposition on how to unify the different theories that have guided both previous research into generics and my experimental approaches. Lastly, Section 7.4 provides some closing remarks and directions for future research.

7.2 Disentangling various types of cues

The guiding motivation for this research at a multidisciplinary intersection is that subtle variations in morphosyntax affect how we conceptualise category structures and the relationship categories have to their properties. Even keeping this question narrow and focused, the results showed the relevance of many other confounding variables. Throughout the discussion I have emphasised that conceptualisation is affected not only subtle morphosyntactic cues (although included in this list for completeness' sake) but also by:

1. morphosyntactic cues
2. (presumed) knowledge about the kind-property connection
3. the amount of exposure/training to a novel kind
4. (presumed) knowledge about the (superordinate) kind
5. visual cues
6. auditory cues

Unlike the predictions based on their semantic representations, the studies have not been able to show that morphosyntactic cues (point 1) can reliably and significantly affect our conceptual system's expectations as to the connection type between kinds and their properties. One of the interactions of wording and property type was observable in the *pseudoword* studies. However, in relative terms morphosyntactic variation does not seem to provide the strongest guidance for the conceptual system and is easily overwritten by other variables. Especially in a series of studies targeting the comprehension of generic subjects, with various types of stimuli, these interactions were either only rarely found, or were found as a function of training (or exposure). Notably, without a training period for the pseudoword studies, the interactions found were opposite to those hypothesised.

It is difficult to assess whether participants additionally and ordinarily rely on their **presumed** category knowledge. If, even in pseudoword studies, they suppose that the novel categories are about e.g. minerals or artefacts or animals, this will influence their judgements in the learning studies (with reference to point 2). This is independent of the fact that participants in these types of studies were not told whether the novel things on the island were animate or inanimate, or any other information about them. Thus, morphosyntax seems to only be a subtle cue, at least in the absence of other linguistic cues, such as prosody.

Analogous to first language acquisition, one might assume that a higher exposure in the training phase should evoke greater effects (point 3). The *pseudoword* studies raised tentative support for this assumption and showed that a medium and an extended amount of training showed some differences between the generic subjects. This leads to further questions about a minimal amount of necessary exposure counterbalanced with exposure that is not so large that it poses a cognitive strain on participants in an experimental setting.

The proposition that morphosyntax is a subtle cue became obvious when analysing the picture book studies, which demonstrated that the role of linguistic cues may be overwritten by a combination of cues. Without further research that specifically isolates these individual factors, it is difficult to judge which one of these are the more influential driving factors in extending less importance to the morphosyntactic form of the subject. Knowing the category type of category a novel animal (at least in broad strokes as a mammal, water-dwelling creature, or a reptile) opens up the space of hypothesis about which features are expected to be prevalent in the instances (with reference to point 4). Thus, any preconceived expectations as to the nature of the basic kind or subkind can affect expectations about category membership and category-property links more strongly than variation in morphosyntax.

Perhaps less surprisingly, the picture books provided extensive visual cues (point 5). The pictures themselves are unlikely to contribute independent information, but rather make more explicit other assumptions about the presumed knowledge of kind membership or kind-property connections, with point 5 thus enhancing effects from points 2 and 4.

The role of auditory cues (referenced in point 6) raises important questions for the phonology-semantics interface. Contrasting intonations to convey information structurally relevant cues such as FOCUS or TOPIC have influenced the creation of the auditory stimuli for studies with child participants. However, it is unclear to what extent auditory cues, without visual information and knowledge about the (superordinate) kind, would affect the importance of morphosyntax. Future research can address this interface, e.g. by adding auditory cues to the *pseudoword* study. If there is an interaction between the expectations raised from FOCUS intonation and the expectations raised by the semantics of the generic subjects, this could enhance the observed distinct behavioural responses, possibly even in a 'no-training' or 'extensive training' condition. For instance, phonologically emphasising a generic subject while using a DS subject could disambiguate a bimodal distribution of responses, as was observed in the *zarpies* studies.

After considering these additional variables and their contribution to generic interpretations, I conclude that while the role of morphosyntax is influential in learning about novel kinds, the human conceptual system does not rely on it exclusively. The role of morphosyntax seems enhanced in the absence of other world knowledge and cues that enable learners to better understanding the organisation of concepts. We have seen that the acceptability of generic subjects is limited by the types of connections they express, such that the DS in particular can only occur in restricted contexts. At the same time, other generic subjects are acceptable in a variety of contexts, since their semantics allow for a wider range of properties to be predicated of their subjects. Thus, we can continue to assume the relevance of morphosyntax in expressing generic knowledge, while also allowing for the contribution of these other confounding factors.

After discussing how these other variables can either enhance or suppress the role of morphosyntax in conceptualisation processes, the next section will outline a draft of how conceptualisation could rely on multiple routes. Having argued that conceptualisation in general depends on a multitude of factors, I will consider the two main routes and how they each might be best suited to the acquisition of concepts, depending on the type of concept and the type of property.

7.3 The proposition: Multiple routes to conceptualisation

I have outlined a number of factors that affect conceptualisation alongside morphosyntactic variation of the subject. This final section ties up some loose ends from the discussion about the underlying nature of categories, as introduced in Chapter 2 and referred to in the discussion of the reverse match-to-sample paradigm in Chapter 6. These discussions aimed at disentangling exposure-based theories from knowledge-driven theories. The experiments on how linguistic variables affect our understanding of novel concepts question link to further debate: Can these findings illuminate whether the observed behavioural differences are due to the fact that different types of properties depend on different types of underlying conceptual structure?

One main reason to suggest multiple routes to conceptualisation is based on the assumption that different types of categories seem to raise expectations about their category structure that are distinct from others, evidenced by our a priori expectations of how a new property will be related to it. To wit, the shape of mathematical objects is expected to play a defining role in their conceptual make-up, yet their colour is not. By contrast, the colour of human blood is expected to be a defining property of the kind HUMAN BLOOD while other properties, e.g. the exact diameter of the veins in which it runs is not. We might then stipulate that the epistemological make-up of categories depends on the intersection of the category type and the property type.

In general, PCs seem more likely to rely on theory theories, or mini-theories as introduced in Section 6.2. By contrast, properties that are true of the category merely by statistical prevalence could be acquired by (repeated) exposure. I have introduced two types of theories that discuss this type of category acquisition: exemplar theory and prototype theory. I have argued that these are two sides of the same coin and that their conceptualisation can be distinguished based on the amount of exposure someone has to members of the kind. I have also commented that combining theory theories with exposure-based theories is not a novel perspective, as it has already been suggested by Osherson & Smith (1981) and Landau (1982) (among others).

One account that proposes multiple routes to the acquisition of concepts and their structures would be to assume that for categories with high essentialism, e.g. animal categories, we rely on the mini-theory component in cases where an underlying biological mechanism is assumed. By contrast, we resort to statistical learning if the predicated properties do not rely on underlying biological mechanisms. This contrast is exemplified by comparing the pattern of an animal's fur with arbitrary properties of animals, similar to the stimuli design for the *kevtas* studies:

(1) a. Tigers are striped.
 b. Tigers play with little bells.

From a cognitive perspective, proposing multiple routes to conceptualisation is not necessarily as resource-intensive as it might initially appear. Returning to the idea of overhypotheses, it is cognitively efficient to rely on underlying mechanisms that explain similarities across a wide range of seemingly vastly different kinds. Thus, it would not be necessary to decide which route to conceptualisation to take on a case-by-case basis, but rather to rely on once-learned means of how different types of properties are normally related to a kind.

From a semantic perspective, by stipulating multiple routes, we can continue to rely on the kind-denoting nature of nominals. This would lead us to expect the prevalence of certain properties within a concept, unless there is reason to assume something is wrong with the kind. When this mechanism fails, e.g. because this kind does not normally possess this type of property, we may rely on exposure and statistical learning.

In the absence of category knowledge, which is easily conveyed through visual cues, we rely heavily on the linguistic form of the generic utterance. The differences, which are based on semantic representations, are enhanced when linguistic cues are the only ones that are available. Interestingly, even limited exposure to less prominent forms of generics subjects during language acquisition is sufficient to allow for these distinctions to develop. We have seen that the English IS and DS are interpreted generically in childhood. Recall, too, that there is a developmental trajectory, such that we can continue to stipulate that the role of exposure is crucial to understanding fine-grained linguistic differences.

Future research should address the idea of acquiring generics via multiple routes by exploring the role of training and exposure with different types of kinds and properties. One major question that is currently left unanswered concerns which conceptualisation route striking properties fall into (although Lazaridou-Chatzigoga et al., 2019's recent work on acquisition of striking properties for novel kinds with child participants has made some excellent points about the graded prevalence expectations that different types of generics raise, such as majority characteristic vs minority characteristic vs striking properties).

To join the findings in the literature and the experimental results in this thesis, I propose that the human conceptual system relies on at least these two routes in concept acquisition, especially when learning about properties that are related to the (novel) kinds in distinct ways. This is ultimately reflected in their different semantics, which are in turn due to their distinct morphosyntactic properties. As the previous sections have shown, the acquisition process is not expected to be linear. Rather, multiple types of cues, in addition to linguistic cues, can modulate expectations of category type and structure. Generic subjects that can be used to express either PCs or SCs, such as the English BP, are expected to be used at a high frequency initially, until sufficient data is available. This is in line with the experimental findings in Chapter 3. To illustrate, learning about a novel kind in the pseudoword paradigm did not allow participants to access superordinate category knowledge, such as the knowledge that they have about animal kinds. In these cases, the BP was typically preferred. This could be explained through a multiple routes account, in which the stipulation of a mini-theory was not necessary. If, with additional exposure in the training phase, the category-property link becomes apparent as more likely to be principled, with an underlying theory to account for this connection, then other subject forms such as the English DS might be preferred (although not necessarily at the expense of the co-occurrence of the BP, as these principally connected properties are normally still statistically prevalent).

7.4 Closing remarks

This thesis has investigated the role of subtle variation in the morphosyntax of generic subjects on the human conceptual system. Through careful manipulation of experimental variables, this thesis has argued that generic subjects are not one homogenous group. Rather, they each convey distinct semantic properties. I tested whether these can result in distinct behavioural responses and found that the results were mixed. Under consideration of the amount of exposure, the type of label (or lack thereof), or presence of world knowledge - among other variables - different generic subject evoke distinct behavioural responses. Yet, the exact contribution of morphosyntax remains to be determined. Within this context, I have demonstrated that the role of morphosyntax in generic subjects is subtle. Its effects are pronounced in the absence of other cues, yet when visual cues and category knowledge are available, the semantic differences between generic subjects are less striking.

Future research should continue to focus on establishing the role of morphosyntactic cues and the conditions under which they exert influence on the conceptualisation of novel kinds. To do so, multiple routes could be taken. In an effort to increase comparability between the findings of the *pseudoword* studies and the picture book studies, a new condition could be introduced to the *pseudoword* studies in which participants are exposed to visual stimuli. This enables an assessment of the extent to which morphosyntactic variations is relied upon when either relying exclusively on pseudowords or both on pseudowords and visual cues. The nature of the visual cues can be highly varied, including animal kinds as in the other studies presented in this thesis as well as a version of the *pseudoword* studies with artefacts or natural kinds such as minerals. In this case, participants then have visual information which allows them to make inferences as to the category type. Crucially, they do not know what the novel predicated property refers to as this remains a pseudoword.

Similarly, an extension of the *kevtas* studies is possible where a new condition is introduced in which participants do not see visual images of the novel animals. Given that the predicated properties still allow to draw limited inferences as to the nature of the novel animal (e.g. *has stripes* or *has two humps*), participants in this extension have different cues available, i.e. category knowledge without visual support. Contrasting with the proposed extension for the *pseudoword* studies, participants can make further inferences as to the connection type since the predicated property is not a pseudoword.

An important follow-up experiment concerns the relationship between amount of necessary exposure and connection type. If the assumption of a dual route to conceptualisation is correct, then we should see that with little exposure, properties that are assumed to be principally connected to a kind (based on previous or assumed knowledge about the kind and how properties of this type are normally connected to it), should easily be generalised. By contrast, properties that are assumed to be statistically connected to a kind should need more exposure to be generalised. While the varying amounts of training only rendered limited effects in the pseudoword studies in Chapter 3, we might observe more distinct judgements by varying the amount of training and also providing information about the category, e.g. in terms of visual stimuli.

In summary, I have illustrated that in order to better understand genericity, we need to take seriously the claim that each generic subject has distinct semantic contributions. These differences can be seen behaviourally in experiments that target, among others, our expectations about category-property links, connection type of the property, and category membership. By introducing the full set of generic subjects of a language into the experimental stimuli, the fine-grained distinctions that we assume in quantified statements due to their overt quantification, can be investigated in generic subjects with covert and overt morphosyntactic differences. This thesis has provided empirical data to linguistic intuitions based on acceptability judgements and underlying semantic work.

Bibliography

Acquaviva, P. (2006). Goidelic inherent plurals and the morphosemantics of number. *Lingua*, *116*(11), 1860–1887.

Acquaviva, P. (2008). *Lexical plurals: A morphosyntactic approach*. Oxford: Oxford University Press.

Aitchison, J. (2012). *Words in the Mind: An Introduction to the Mental Lexicon* (4th ed.). Wiley-Blackwell.

Alexiadou, A. (2014). *Multiple Determiners and the Structure of DPs (Linguistik Aktuell/Linguistics Today, Vol. 211)*. Amsterdam: John Benjamins.

Alexiadou, A., Haegeman, L., & Stavrou, M. (2007). *Noun Phrase in the Generative Perspective*. de Gruyter.

Allerton, D. J. & Cruttenden, A. (1979). Three reasons for accenting a definite subject. *Journal of Linguistics*, *15*(1), 49–53.

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of Social Theories: The Role of Explanation in the Persistence of Discredited Information. *Journal of Personality and Social Psychology*, *39*(6), 1037–1049.

Arka, I. W. & Dalrymple, M. (2016). Number and plural semantics: Empirical evidence from Marori. *NUSA: Linguistic Studies of Languages In and Around Indonesia*, *60*(30), 89–106.

Asher, N. & Morreau, M. (1995). What some generic sentences mean. In G. N. Carlson & F. J. Pelletier (Eds.), *The Generic Book* chapter 7, (pp. 300–338). Chicago: The University of Chicago Press.

Baldwin, D. A. (1992). Clarifying the role of shape in children's taxonomic assumption. *Journal of Experimental Psychology*, *54*, 392–416.

Barker, C. (1992). Group Terms in English: Representing Groups as Atoms. *Journal of Semantics*, *9*(1), 69–93.

Barsalou, L. W. (2000). Concepts: Structure. In A. Kazdin (Ed.), *Encyclopedia of Psychology. Vol. 2* (pp. 245–248). Oxford: Oxford University Press.

Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, *8*(4), 243–257.

Bierwisch, M. (1970). On Classifying Semantic Features. In M. Bierwisch & K. E. Heidolph (Eds.), *Progress in Linguistics: A Collection of Papers* (pp. 27–50). The Hague: Mouton.

Bird, A. & Tobin, E. (2018). Natural Kinds. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring2018 ed.). Metaphysics Research Lab, Stanford University.

Blair, R. & Vermont-Salas, R. (1967). *Spoken Yucatec Maya.* PhD thesis, University of Chicago.

Blok, S., Newman, G., & Rips, L. J. (2005). Individuals and Their Concepts. In W.-k. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin.* (pp. 127–149). Washington, D. C.: American Psychological Association.

Bloom, P. (1999). Theories of word learning: Rationalist alternatives to associationism. In T. K. Bhatia & W. C. Ritchie (Eds.), *Handbook of language acquisition.* New York: Academic Press.

Booth, A. E. & Waxman, S. R. (2006). Déjà vu all over again: Rerevisiting the conceptual status of early word learning: Comment on Smith and Samuelson. *Developmental Psychology*, *42*, 1344–1346.

Booth, A. E., Waxman, S. R., & Huang, Y. T. (2005). Conceptual information permeates word learning in infancy. *Developmental Psychology*, *41*(3), 491–505.

Borer, H. (2005). *Structuring Sense: Volume 2: The Normal Course of Events.* Oxford: Oxford University Press.

Borer, H. (2009). Syntax for Kinds. In *Mass/Count Workshop.* University of Toronto.

Borer, H. & Swart, D. (2011). Bare Nouns, Kindly. In *Weak Referentiality Workshop.* Utrecht Institute of Linguistics.

Borik, O. (2016). Constraints on the position and interpretation of bare singular indefinites in Russian. *Linguistica*, *56*(1), 9–23.

Borik, O. & Espinal, M. T. (2012). On definite kinds. *Recherches linguistiques de Vincennes*, *41*, 123–146.

Borik, O. & Espinal, M. T. (2015). Reference to kinds and to other generic expressions in Spanish: definiteness and number. *The Linguistic Review*, (1), 1–58.

Bowdle, B. F. & Ward, G. (1995). Generic Demonstratives. In *Proceedings of the Twenty-First Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Historical Issues in Sociolinguistics/Social Issues in Historical Linguistics*, (pp. 32–43).

Brandone, A. C., Cimpian, A., Leslie, S.-J., & Gelman, S. A. (2012). Do Lions Have Manes? For Children, Generics Are About Kinds Rather Than Quantities. *Child Development*, *83*(2), 423–433.

Brennan, V. M. (1993). *Root and Epistemic Modal Auxiliary Verbs.* Doctoral dissertation, Amherst: University of Massachussetts.

Brent, M. & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*, 31–44.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Bruner, J. S., Olver, R. R., & Greenfield, P. M. (1966). *Studies in cognitive growth.* New York: Wiley.

Burton-Roberts, N. (1976). On the Generic Indefinite Article. *Linguistic Society of America*, *52*(2), 427–448.

Butler, L. K. (2012). Crosslinguistic and experimental evidence for non-number plurals. *Linguistic Variation*, *1*, 27–56.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Carey, S. (1988). Conceptual Differences Between Children and Adults. *Mind & Language*, *3*(3), 167–181.

Carey, S. (2009). *The Origin of Concepts.* Oxford: Oxford University Press.

Carlson, G. N. (1977). *Reference to Kinds in English.* Doctoral dissertation, University of Massachusetts/Amherst.

Carlson, G. N. (1980). The Semantic Composition of English Generic Sentences. Technical report, University of Iowa.

Carlson, G. N. (1989). The semantic composition of English generic sentences. In G. Chierchia, B. Partee, & R. Turner (Eds.), *Properties, types and meaning* (pp. 167–192). Kluwer.

Carlson, G. N. (2010). Generics and Concepts. In F. J. Pelletier (Ed.), *Kinds, Things, and Stuff. Mass Terms and Generics* chapter 2, (pp. 16–35). Oxford: Oxford University Press.

Carlson, G. N. & Pelletier, F. J. (1995). *The Generic Book.* Chicago: The University of Chicago Press.

Carroll, L. (1871). *Through the Looking-Glass.* Macmillan.

Carston, R. (2012). Word meaning and concept expressed. *Linguistic Review*, *29*(4), 607–623.

Cavedon, L. & Glasbey, S. (1994). Outline of an Information-Flow Model of Generics. *Acta Linguistica Hungarica*, *42*.

Cheng, L. L. S., Doetjes, J., & Sybesma, R. (2008). How universal is the Universal Grinder? *Linguistics in the NetherlandsLinguistics in the Netherlands. AVT Publications*, *25*(January 2008), 50–62.

Chierchia, G. (1995). Individual Level Predicates as Inherent Generics. In G. N. Carlson & F. J. Pelletier (Eds.), *The Generic Book* (pp. 176–224). Chicago: Chicago University Press.

Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, *6*, 339–405.

Cimpian, A., Brandone, A. C., & Gelman, S. A. (2010). Generic Statements Require Little Evidence for Acceptance but Have Powerful Implications. *Cognitive Science*, *34*(8), 1452–1482.

Cimpian, A. & Cadena, C. (2010). Why are dunkels sticky? Preschoolers infer functionality and intentional creation for artifact properties learned from generic language. *Cognition*, *117*, 62–68.

Cimpian, A. & Markman, E. M. (2009). Information learned from generic language becomes central to children's biological concepts: Evidence from their open-ended explanations. *Cognition*, *113*(1), 14–25.

Cimpian, A., Meltzer, T. J., & Markman, E. M. (2011). Preschoolers' use of morphosyntactic cues to identify generic sentences: Indefinite singular noun phrases, tense, and aspect. *Child Development*, *82*(5), 1561–1578.

Clark, E. V. & Clark, H. H. (1979). When nouns surface as verbs. *Language*, *55*, 767–811.

Cohen, A. (2005). More than bare existence: An implicature of existential bare plurals. *Journal of Semantics*, *23*, 389–400.

Coppock, E. & Beaver, D. (2015). Definiteness and determinacy. *Linguistics and Philosophy*, *38*(5), 377–435.

Craig, C. G. (1986). Jacaltec noun classifiers. A study in grammaticalization. *Lingua*, *70*, 241–284.

Dahl, Ö. (1975). On Generics. In E. Keenan (Ed.), *Formal Semantics of Natural Language* (pp. 99–111). Cambridge: Cambridge University Press.

Davidson, N. S. & Gelman, S. A. (1990). Inductions from novel categories: The role of language and conceptual structure. *Cognitive Development*, *5*, 151–176.

Davis, B. & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, *38*, 1199–1211.

Dayal, V. (1992). The Singular-Plural Distinction in Hindi Generics.

Dayal, V. (1999). Bare NP's, Reference to Kinds, and Incorporation. In Matthews, T. & Strolovitch, D. (Eds.), *Proceedings of SALT IX*, (pp. 34–51)., Ithaca, NY. Cornell University.

Dayal, V. (2004). Number marking and (in) definiteness in kind terms. *Linguistics and Philosophy*, *27*(4), 393–450.

Declerk, R. (1986). The Manifold Interpretations of Generic Sentences. *Lingua*, *68*, 149–188.

Declerk, R. (1991). The origins of genericity. *Linguistics*, *29*(1), 79–102.

Dewar, K. M. & Xu, F. (2010). Induction, Overhypothesis, and the Origin of Abstract Knowledge: Evidence From 9-Month-Old Infants. *Psychological Science*, *21*(12), 1871–1877.

Diesendruck, G., Markson, L., & Bloom, P. (2003). Children's reliance on creator's intent when naming artifacts. *Psychological Science*, *14*, 164–168.

Diesing, M. (1992). *Indefinites*. Cambridge, MA: MIT Press.

Diessel, H. (2009). Corpus Linguistics and Language Acquisition. In Lüdeling, A. & Kytö, M. (Eds.), *Corpus Linguistics. An International Handbook*, (pp. 1197–1212)., Berlin. Mouton de Gruyter.

Ferrand, L. (2010). The French Lexicon Project.

Filipovic, J. (2016). Go Ahead, Play the Woman Card.

Fodor, J. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, J. (1981). The present status of the innateness hypothesis. In J. Fodor (Ed.), *RePresentations: Philosophical Essays on the Foundations of Cognitive Science.* (pp. 257–316). Cambridge, MA: MIT Press.

Fodor, J. (1998). *Concepts: Where Cognitive Science Went Wrong*. Cambridge, MA: MIT Press.

Fodor, J. (2003). *Hume Variations*. Oxford: Oxford University Press.

Fodor, J. & Lepore, E. (1996). The red herring and the pet fish: why concepts still can't be prototypes. *Cognition*, *58*, 253–270.

Frege, G. (1948). Sense and reference. *The Philosophical Review*, *57*(3), 209–230.

Fuellenbach, K. (2017). *Genericity and the human conceptual system: On the morphosyntax of generic subjects*. M.phil. thesis, University of Oxford.

Fuellenbach, K. & Gelman, S. A. (2019). Generic and non-generic interpretations for singular and plural subjects. Experimental studies on genericity. In van Alem, A., De Sisto, M., Kerr, E. J., & Wall, J. (Eds.), *Proceedings of the 27th Conference of the Student Organization of Linguistics in Europe*, (pp. 38–58). Leiden University Centre for Linguistics.

Fuellenbach, K., Gelman, S. A., Husband, E. M., & Cuneo, N. (2019). Generalising properties based on the morphosyntax of the subject : generic and non-generic interpretations. Technical report, 25th Architectures and Mechanisms of Language Processing, Moscow.

Fulkerson, A. L. & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition*, *105*, 218–228.

Geeraerts, D. (2008). Prospects and problems of prototype theory. In *Cognitive Linguistics: Basic Readings* chapter 4, (pp. 141–165). Berlin: Mouton de Gruyter.

Gelman, S. A. (2003). *The essential child: origins of essentialism in everyday thought*. Oxford series in cognitive development. Oxford ; New York: Oxford University Press.

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, *8*(9), 404–409.

Gelman, S. A., Chesnick, R. J., & Waxman, S. R. (2005). Mother-child conversations about pictures and objects: Referring to categories and individuals. *Child Development*, *76*(6), 1129–1143.

Gelman, S. A. & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in two-year-old children. *Developmental Psychology*, *26*, 796–804.

Gelman, S. A., Coley, J. D., Rosengren, K. S., Hartman, E., Pappas, A., & Keil, F. C. (1998). *Beyond Labeling: The Role of Maternal Input in the Acquisition of Richly Structured Categories*, volume 63. WileySociety for Research in Child Development.

Gelman, S. A., Heyman, G. D., & Legare, C. H. (2007). Developmental changes in the coherence of essentialist beliefs about psychological characteristics. *Child Development*, *78*, 757–774.

Gelman, S. A. & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*(3), 183–209.

Gelman, S. A. & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, *59*, 876–887.

Gelman, S. A. & Raman, L. (2003). Preschool Children Use Linguistic Form Class and Pragmatic Cues to Interpret Generics. *Child Development*, *74*(1), 308–325.

Gelman, S. A., Sánchez Tapia, I., & Leslie, S.-J. (2016). Memory for generic and quantified sentences in Spanish-speaking children and adults. *Journal of Child Language*, *43*, 1231–1244.

Gelman, S. A. & Tardif, T. (1998). A cross-linguistic comparison of generic noun phrases in English and Mandarin. *Cognition*, *66*(3), 215–248.

Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology*, *61*(3), 273–301.

Gelman, S. A., Ware, E. A., Manczak, E. M., & Graham, S. A. (2013). Children's Sensitivity to the Knowledge Expressed in Pedagogical and Non-Pedagogical Contexts. *Developmental Psychology*, *49*(3), 491–504.

Goel, V., Gold, B., Kapur, S., & Houle, S. (1997). The seats of reason? An imaging study of deductive and inductive reasoning. *NeuroReport*, *8*, 1305–1310.

Goldin-Meadow, S., Gelman, S. A., & Mylander, C. (2005). Expressing generic concepts with and without a language model. *Cognition*, *96*, 109–126.

Goodman, N. (1983). *Fact, Fiction, and Forecast* (4th ed.). Cambridge, MA: Harvard University Press.

Gopnik, A. & Meltzoff, A. N. (1997). *Words, Thoughts, and Theories.* Cambridge, MA: MIT Press.

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive references. *Child Development*, *75*, 409–427.

Greenberg, J. H. (1963). *Universals of Language.* London: MIT Press.

Greenberg, Y. (2003). *Manifestations of Genericity.* New York: Routledge.

Greenberg, Y. (2012). Genericity and (Non-)Accidentalness. *Recherches linguistiques de Vincennes*, *41*, 163–190.

Griffiths, P. E. (1997). *What Emotions Really Are: The Problem of Psychological Categories.* Chicago: University of Chicago Press.

Griffiths, P. E. (2002). Is emotion a natural kind?

Gunkel, L., Murelli, A., Schlotthauer, S., Wiese, B., & Zifonun, G. (2017). *Grammatik des Deutschen im europäischen Vergleich: Das Nominal [Band 1].* de Gruyter.

Harris, P. L. & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child Development*, *77*(3), 505–524.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*, 113–127.

Hawkins, J. A. (2004). *More on Form Minimization.* Oxford: Oxford University Press.

Heim, I. (1982). *The Semantics of Definite and Infedinite Noun Phrases.* Phd thesis, University of Massachusetts.

Heim, I. & Kratzer, A. (1998). *Semantics in Generative Grammar.* Blackwell.

Helweg, H. (1978). *Farm Animals.* New York: Random House, Inc.

Heycock, C. & Zamparelli, R. (2005). Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural Language Semantics*, *13*(3), 201–270.

Hollander, M. A. (2007). *Learning and Interpreting Words for Kinds: Adults' and children's understanding of generic language.* Phd thesis, University of Michigan.

Hollander, M. A., Gelman, S. A., & Raman, L. (2009). Generic language and judgements about category membership: Can generics highlight properties as central? *Language and Cognitive Processes*, *24*(4), 481–505.

Hume, D. (1739). *A Treatise of Human Nature.* Oxford: Oxford University Press.

Hume, D. (1748). *An enquiry concerning human understanding.* Indianapolis, IN: Hackett.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development*, *9*, 45–75.

Jaswal, V. K. & Markman, E. M. (2002). Children's acceptance and use of unexpected category labels to draw non-obvious inferences. In Gray, W. D. & Schunn, C. (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society*, (pp. 500–505). Erlbaum.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA: MIT Press.

Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences*, *7*(8), 368–373.

Kelemen, D. & Bloom, P. (1994). Domain-specific knowledge in simple categorization tasks. *Psychonomic Bulletin & Review*, *1*(3), 390–395.

Kelly, G. A. (1955). *The psychology of personal constructs. Volume I. A theory of personality.* New York: Norton.

Kemler Nelson, D. G., Frankenfield, A., Morris, C., & Blair, E. (2000). Young children's use of functional information to recognize artifacts: Three factors that matter. *Cognition*, *77*, 133–168.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models (BAYESIAN SPECIAL SECTION). *Developmental Science*, *10*(3), 307–321.

Khemlani, S., Leslie, S.-J., & Glucksberg, S. (2012). Inferences about members of kinds: The generics hypothesis. *Language and Cognitive Processes*, *27*(6), 887–900.

Khemlani, S., Leslie, S.-J., Glucksberg, S., & Fernandez, P. R. (2007). Do ducks lay eggs? How people interpret generic assertions. In *Proceedings of the 29th annual conference of the cognitive science society*, (pp. 395–401).

Kirkpatrick, J. (2020). Generic Knowledge. In *Talk given at the Cambridge Generics Workshop*, Cambridge.

Kratzer, A. (1995). Stage-level and individual-level predicates. In G. N. Carlson & F. J. Pelletier (Eds.), *The Generic Book* chapter 2, (pp. 125–175). Chicago: The University of Chicago Press.

Kratzer, A. & Selkirk, E. (2007). Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review*, *24*, 93–135.

Krifka, M. (2003). Bare NPs : Kind-referring, Indefinites, Both, or Neither? In Young, R. & Zhou, Y. (Eds.), *Proceedings of SALT XIII*, (pp. 180–203)., Ithaca, NY. Cornell University.

Krifka, M. (2017). Generic Sentences as Quantifications over Samples.

Krifka, M., Pelletier, F. J., Carlson, G. N., ter Meulen, A., Link, G., & Chierchia, G. (1995). Genericity: An Introduction. In G. N. Carlson & F. J. Pelletier (Eds.), *The Generic Book* (pp. 1–124). The University of Chicago Press.

Kwapiszewski, A. & Fuellenbach, K. (2020). Reference to Kinds and Subkinds in Polish. In *Advances in Formal Slavic Linguistics 2018*. Language Science Press.

Landau, B. (1982). Will the Real Grandmother Please Stand Up? The Psychological Reality of Dual Meaning Representations. *Journal of Psycholinguistic Research*, *11*(1), 47–62.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321.

Lawler, J. (1973). Studies in English Generics. Technical Report 1, University of Michigan.

Lazaridou-Chatzigoga, D., Katsos, N., & Stockall, L. (2015). Genericity is Easy? Formal and Experimental Perspectives. *Ratio*, *28*(4), 470–494.

Lazaridou-Chatzigoga, D., Katsos, N., & Stockall, L. (2019). Generalizing about striking properties: Do glippets love to play with fire? *Frontiers in Psychology*, *10*.

Lazaridou-Chatzigoga, D. & Stockall, L. (2013). Genericity, exceptions and domain restriction: experimental evidence from comparison with universals. In Chemla, E., Homer, V., & Winterstein, G. (Eds.), *Proceedings of SuB17, École Normale Supérieure, Paris*, (pp. 325–343)., Paris.

Lazaridou-Chatzigoga, D., Stockall, L., & Katsos, N. (2019). Contextualising generic and universal generalisations: quantifier domain restriction and the generic overgeneralisation effect. *Journal of Semantics*, *36*, 617–664.

Lee, C. (1996). Generic Sentences are Topic Constructions. In T. Fretheim & J. K. Gundel (Eds.), *Reference and Referent Accessibility. Pragmatics & Beyond New Series. Vol. 38* (pp. 213–222). John Benjamins.

Leonard, N. & Rips, L. J. (2015). Identifying and counting objects: The role of sortal concepts. *Cognition*, *145*, 89–103.

Leslie, S.-J. (2008). Generics: Cognition and Acquisition. *Philosophical Review*, *117*(1), 1–47.

Leslie, S.-J. & Gelman, S. A. (2012). Quantified Statements are Recalled as Generics: Evidence from Preschool Children and Adults. *Cognitive Psychology*, *64*(3), 186–214.

Leslie, S.-J., Khemlani, S., & Glucksberg, S. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language*, *65*(1), 15–31.

Leslie, S.-J. & Lerner, A. (2016). Generic Generalizations. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter2016 ed.). Metaphysics Research Lab, Stanford University.

Leu, T. (2015). *The Architecture of Determiners.* Oxford: Oxford University Press.

Lewis, D. (1975). Adverbs of Quantification. In E. L. Keenan (Ed.), *Formal Semantics of Natural Language* (pp. 3–15). Cambridge: Cambridge University Press.

Liebesman, D. (2011). Simple Generics. *Noûs*, *45*(3), 409–442.

Link, G. (1983). The logical analysis of plural and mass terms: A lattice theoretical approach. *Meaning, use, and interpretation of language*, (9), 302–323.

Löbner, S. (2002). *Understanding Semantics*. London: Arnold.

Locke, J. (1690). *An Essay Concerning Human Understanding*. New York: Oxford University Press.

Longobardi, G. (2000). Remarks and replies: "Postverbal" subjects and the mapping hypothesis. *Linguistic Inquiry*, *31*(4), 691–702.

Longobardi, G. (2001). The Structure of DPs: Some principles, parameters, and problems. In M. Baltin & C. Collins (Eds.), *The Handbook of Contemporary Syntactic Theory* (pp. 562–603). Oxford: Blackwell.

Ludlow, P. (2008). Cheap Contextualism. *Philosophical Issues*, *18*, 104–129.

Machery, E. (2009). *Doing Without Concepts*. New York: Oxford University Press.

MacWhinney, B. C. M. U. (2000). The CHILDES Project: Tools for Analyzing Talk. Technical report, Carnegie Mellon University, Mahwah, NJ.

Mancuso, J. C. (1996). Constructionism, personal construct psychology and narrative psychology. *Theory & Psychology*, *6*, 47–70.

Marchak, K. & Gelman, S. A. (2018). How do we learn overhypotheses about the characteristic properties of different kinds of animals? In *Poster presented at the 44th annual meeting of the Society for Philosophy and Psychology*, (pp.2̃)., Ann Arbor, MI. University of Michigan.

Mari, A., Beyssade, C., & Del Prete, F. (2013). *Genericity*. Oxford: Oxford University Press.

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua*, *92*(1-4), 199–227.

McCloskey, M. E. & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory and Cognition*, *6*, 462–472.

McNally, L. (2017). Kinds, descriptions of kinds, concepts, and distributions. In K. Balogh & W. Petersen (Eds.), *Bridging Formal and Conceptual Semantics. Selected papers of BRIDGE-14* (pp. 39–61). Düsseldorf: DUP.

McNally, L. & Boleda, G. (2004). Relational adjectives as properties of kinds. In O. Bonami & P. Cabredo Hofherr (Eds.), *Empirical Issues in Formal Syntax and Semantics*, volume 5 (pp. 179–196). Colloque de Syntaxe et Semantique a Paris.

McNaught, H. (1978). *The Truck Book*. New York: Random House, Inc.

McNorgan, C., Kotack, R. A., Meehan, D. C., & McRae, K. (2007). Feature-feature causal relations and statistical co-occurrences in object concepts. *Memory and Cognition*, *35*(3), 418–431.

Medin, D. L. & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

Medin, D. L. & Smith, E. E. (1981). Strategies and classification learning. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(4), 241–253.

Mervis, C. B. & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*, 89–115.

Meyer, M. & Gelman, S. A. (2016). Generic Reference is Less Marked Than Specific Reference in Children's Gestures. *Journal of Nonverbal Behavior*, *40*(1), 65–79.

Meyer, M., Gelman, S. A., & Stilwell, S. (2011). Generics are a cognitive default: evidence from sentence processing. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, (pp. 913 – 918).

Mueller-Reichau, O. (2013). Sorting the world: On the relevance of the kind/object-distinction to referential semantics. In G. Grewendorf, W. Hinzen, H. Kamp, & H. Weiß (Eds.), *Linguistics and Philosophy (4)* (pp. 165). Berlin, Boston: de Gruyter.

Murphy, G. L. (1993). Theories and concept formation. In I. van Mechelen, J. A. Hampton, R. Michalski, & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 173–200). New York: Academic Press.

Murphy, G. L. (2002). *The Big Book of Concepts.* Cambridge, MA: MIT Press.

Murphy, G. L. (2005). The study of concepts inside and outside of the laboratory: Medin versus Medin. In W.-k. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin.* (pp. 179–195). Washington, D. C.: American Psychological Association.

Murphy, G. L. & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*(3), 289–316.

Nguyen, A. (2019). The radical account of bare plural generics. *Philosophical Studies*, *1*, 1–29.

Nickel (2016). *Between logic and the world: An integrated theory of generics.* Oxford: Oxford University Press.

Nickel, B. (2012). Dutchmen are Good Sailors: Generics and Gradability. In A. Mari, C. Beyssade, & F. Del Prete (Eds.), *Genericity* chapter 16. Oxford: Oxford University Press.

O'Grady, W. (2005). *How Children Learn Language.* Cambridge: Cambridge University Press.

Oltean, S. (2009). On the denotations of common nouns. *Philologia*, *3*, 263–272.

Osherson, D., Perani, D., Cappa, S., Schnur, T., Grassi, F., & Fazio, F. (1998). Distinct brain loci in deductive versus probabilistic reasoning. *Neuropsychologia*, *36*, 369–376.

Osherson, D. & Smith, E. E. (1981). On the Adequacy of Prototype Theory as a Theory of Concepts. *Cognition*, *9*, 35–58.

Partee, B. (1974). *Montague Grammar*. New York: Academic Press.

Partee, B. (1987). Noun Phrase Interpretation and Type-Shifting Principles. In J. Groenendijk, D. de Jongh, & M. Stokhof (Eds.), *Studies in discourse representation theory and the theory of generalized quantifiers* chapter 5, (pp. 115–141). Amsterdam: Foris.

Pelletier, F. J. (1975). Non-singular reference: some preliminaries. *Philosophia*, *5*(4), 451–465.

Pelletier, F. J. (2010). *Kinds, Things, and Stuff*. Oxford: Oxford University Press.

Picallo, M. C. (2006). Some notes on grammatical gender and l-pronouns. In K. von Heusinger, G. A. Kaiser, & E. Stark (Eds.), *Specificity and the evolution/emergence of nominal determination systems in Romance (Arbeitspaper 119)* (pp. 107–121). Konstanz: University of Konstanz.

Pinker, S. (1994). *The Language Instinct: The New Science of Language and Mind*. London: Penguin.

Pinker, S. (2007). *The Stuff of Thought: Language as a Window into Human Nature*. London: Penguin.

Plunkett, D. (2015). Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy. *Inquiry*, *58*(7-8), 828–874.

Plunkett, D. & Sundell, T. (2019). Metalinguistic negotiation and speaker error. *Inquiry*, *1*, 1–26.

Plunkett, K., Hu, J. F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, *106*(2), 665–681.

Posner, M. I. & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.

Posner, M. I. & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, *83*, 304–308.

Prasada, S. (2000). Acquiring generic knowledge. *Trends in Cognitive Sciences*, *4*(2), 66–72.

Prasada, S. (2010). Conceptual Representation and Some Forms of Genericity. In F. J. Pelletier (Ed.), *Kinds, Things, and Stuff. Mass Terms and Generics* chapter 3, (pp. 36–59). Oxford: Oxford University Press.

Prasada, S. (2016a). An explicit theory of sortal representation and some evidence for it.

Prasada, S. (2016b). Mechanisms for thinking about kinds, instances of kinds, and kinds of kinds. In D. Barner & A. S. Baron (Eds.), *Core knowledge and conceptual change* chapter 12, (pp. 209–224). Oxford: Oxford University Press.

Prasada, S. & Dillingham, E. M. (2006). Principled and statistical connections in common sense conception. *Cognition*, *99*(1), 73–112.

Prasada, S. & Dillingham, E. M. (2009). Representation of principled connections: A window onto the formal aspect of common sense conception. *Cognitive Science*, *33*(3), 401–448.

Prasada, S., Ferenz, K., & Haskell, T. (2002). Conceiving of entities as objects and as stuff. *Cognition*, *83*(2), 141–165.

Prasada, S. & Hall, D. G. (2019). Instance-of-object-kind representations. *Cognition*, *189*, 209–220.

Prasada, S., Khemlani, S., Leslie, S.-J., & Glucksberg, S. (2013). Conceptual distinctions amongst generics. *Cognition*, *126*(3), 405–422.

Quine, W. v. O. (1960). *Word and Object*. Cambridge, MA: MIT Press.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, *3*, 382–407.

Reiter, N. & Frank, A. (2010). Identifying Generic Noun Phrases. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, (pp. 40–49).

Rey, G. (2014). Innate and Learned: Carey, Mad Dog Nativism, and the Poverty of Stimuli and Analogies (Yet Again). *Mind and Language*, *29*(2), 109–132.

Rhemtulla, M. & Xu, F. (2007). Sortal concepts and causal continuity: comment of Rips, Blok, and Newman (2006). *Psychological review*, *114*(4), 1087–1095.

Rhodes, M., Leslie, S.-J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(34), 13526–13531.

Rips, L. J. (2001). Two kinds of reasoning. *Psychological Science*, *12*(2), 129–134.

Rips, L. J. (2011). *Lines of Thought. Central Concepts in Cognitive Psychology.* Oxford: Oxford University Press.

Rips, L. J., Blok, S., & Newman, G. (2006). Tracing the identity of objects. *Psychological Review*, *113*(1), 1–30.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, *12*, 1–20.

Rollins, P. R. (2003). Caregiver contingent comments and subsequent vocabulary comprehension. *Applied Psycholinguistics*, *24*, 221–234.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. LLoyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rosch, E. & Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic Objects in Natural Categories. *Cognitive Psychology*, *8*, 382–439.

Rozenblit, L. & Keil, F. C. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science*, *26*, 521–562.

Russell, B. (1919). *Introduction to Mathematical Philosophy*. London: George Allen and Unwin.

Sallmayer (1865). Der Satz und die Satzzeichen - Volume 1.

Sanches, M. & Slobin, L. (1973). Numeral classifiers and plural marking: An implicational universal. In *Working Papers on Language Universals*, Working Papers on Language Universals, (pp. 1–22)., Stanford, CA. Stanford University.

Sapir, E. (1921). *Language: An Introduction to the Study of Speech*. New York: Harcourt Brace Jovanovich.

Sawyer, K. (1997). *Pretend play as improvisation*. Mahwah, NJ: Erlbaum.

Saxton, M. (2010). *Child language. Acquisition and development*. London: Sage.

Schmitt, C. & Munn, A. (1999). Against the Nominal Mapping Parameter: Bare nouns in Brazilian Portuguese. In *Proceedings of NELS29*, volume 29, (pp. 339–354).

Schwarz, F. (2009). *Two types of definites in natural language*. PhD thesis, University of Massachusetts.

Schwarzschild, R. (1996). *Pluralities*. Dordrecht/Boston/London: Kluwer.

Smith, E. E. & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Smith, L. B. (1989). A model of perceptual classification in children and adults. *Psychological Review*, *96*, 125–144.

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: a dumb attentional mechanism? *Cognition*, *60*, 143–171.

Sprouse, J. & Almeida, D. (2012). Assessing the reliability of textbook data in syntax: Adger's Core Syntax. *Journal of Linguistics*, *48*(3), 609–652.

Sprouse, J. & Almeida, D. (2017). Design sensitivity and statistical power in acceptability judgment experiments. *Glossa: a journal of general linguistics*, *2*(1), 14.

Stam, H. J. (1998). Personal-construct theory and social constructionism: Difference and dialogue. *Journal of Constructivist Psychology*, *11*(3), 187–203.

Sterken, R. K. (2015a). Generics, Content and Cognitive Bias. *Analytic Philosophy*, *56*(1), 75–93.

Sterken, R. K. (2015b). Generics in Context. *Philosophers' Imprint*, *15*(21).

Sterken, R. K. (2015c). Leslie on Generics. *Philosophical Studies*, *172*(9), 2493–2512.

Sterken, R. K. (2020). Generics and Metalinguistic Negotiations. In *Talk given at the Cambridge Generics Workshop*, Cambridge.

Strawson, P. F. (1950). On Referring. *Mind*, *59*(235), 3320–344.

Strickland, K. (2017). Why Men Should Stop Saying #NotAllMen. Immediately.

Sutherland, S. L., Cimpian, A., Leslie, S.-J., & Gelman, S. A. (2015). Memory Errors Reveal a Bias to Spontaneously Generalize to Categories. *Cognitive Science*, *39*(5), 1021–1046.

Tangen, J. M. & Allan, L. G. (2004). Cue interaction and judgments of causality: Contributions of causal and associative processes. *Memory and Cognition*, *32*, 107–124.

Teichman, M. (2015). *Characterizing Kinds. A Semantics for Generic Sentences.* Phd thesis, University of Chicago.

ter Meulen, A. (1986). Generic Information, Conditional Contexts and Constraints. In E. C. Traugott, A. ter Meulen, J. S. Reilly, & C. A. Ferguson (Eds.), *On Conditionals* (pp. 123–146). Cambridge: Cambridge University Press.

Tomasello, M. (2003). *Constructing a language. A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Tomasello, M. & Stahl, D. (2004). Sampling children's spontaneous speech: How much is enough? *Journal of Child Language*, *31*, 101–121.

Topolinjska, Z. (2006). Toward a definition of the generic noun phrase. *Linguistica (Ljubljana, Print)*, *46*, 55–59.

Vendler, Z. (1971). Singular terms. In D. Steinberg & L. Jakobovitz (Eds.), *Semantics: An Interdisciplinary Reader* (pp. 115–133). Cambridge: Cambridge University Press.

von Fintel, K. (1996). Specific Generics.

Vygotsky, L. S. (1962). *Thought and language.* New York: Wiley.

Wągiel, M. (2014). From kinds to objects: prenominal and postnominal adjectives in Polish. In *Complex Visibles Out There: Proceedings of the Olomouc Linguistics Colloquium. Language Use and Linguistic Structure*, (pp. 457–477).

Waxman, S. R. (2003). Links between object categorization and naming: origins and emergence in human infants. In D. H. Rakison & L. M. Oakes (Eds.), *Early Category and Concept Development: Making Sense of the Blooming, Buzzing Confusion* (pp. 213–241). Oxford: Oxford University Press.

Waxman, S. R. & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-month-old infants. *Cognitive Psychology*, *29*, 257–302.

Webelhuth, G. (1989). *Syntactic Saturation Phenomena and the Modern Germanic Languages.* Ph.d. thesis, University of Massachussetts, Amherst.

Weisgerber, L. (1950). Vom Weltbild der deutschen Sprache.

Whorf, B. L. (1956). Language, thought, and reality. In J. Carroll (Ed.), *Selected Writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.

Wiltschko, M. (2008). The syntax of non-inflectional plural marking. *Natural Language and Linguistic Theory*, *26*, 639–694.

Wisniewski, E. J. (1995). Prior Knowledge and Functionally Relevant Features in Concept Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 449–468.

Wittgenstein, L. (1953). *Philosophical Investigations*. New York: Macmillan.

Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, *6*(3), 297–311.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, *85*(3), 223–250.

Zamparelli, R. (2000). *Layers in the Determiner Phrase*. Garland Publishing Inc.

Appendices

A
Corpus Analysis

Contents

A.1 Example table for MLU in CHILDES corpus

Table of #utterances, #morphemes, morphemes utterance (i.e. mean length utterance [MLU]), and Standard Deviation (SD) for files "c1" in Brent & Siskind (2001):

Table A.1: Example of summary table of meta-data per mother-child dyad, broken down by recording

FILE	LANG	CORPUS	CODE	AGE	SEX	ROLE	#UTT	#MOR	MOR/UTT	SD
902	eng	Brent	PAR	.	.	Participant	0	0	N/A	N/A
000917	eng	Brent	MOT	.	female	Mother	1231	4839	3.931	2.565
917	eng	Brent	CHI	0;09.17	female	Child	11	18	1.636	1.367
000930	eng	Brent	MOT	.	female	Mother	950	4621	4.864	3.058
001014	eng	Brent	MOT	.	female	Mother	527	2302	4.368	3.014
1014	eng	Brent	CHI	0;10.14	female	Child	5	23	4.6	2.577
001027	eng	Brent	MOT	.	female	Mother	842	3778	4.487	3.15
1027	eng	Brent	CHI	0;10.27	female	Child	9	16	1.778	2.2
001129	eng	Brent	MOT	.	female	Mother	894	3820	4.273	2.859
1129	eng	Brent	CHI	0;11.29	female	Child	22	28	1.273	0.914
010007	eng	Brent	MOT	.	female	Mother	772	3281	4.25	2.835
10007	eng	Brent	CHI	1;00.07	female	Child	7	7	1	0
010017	eng	Brent	MOT	.	female	Mother	592	2398	4.051	2.975
10017	eng	Brent	CHI	1;00.17	female	Child	29	30	1.034	0.182
010109	eng	Brent	OTH	.	.	Unidentified	3	11	3.667	3.091
10109	eng	Brent	MOT	.	female	Mother	796	3010	3.781	2.904
10109	eng	Brent	CHI	1;01.09	female	Child	26	28	1.077	0.266
010121	eng	Brent	OTH	.	.	Unidentified	22	66	3	1.706
10121	eng	Brent	MOT	.	female	Mother	536	1908	3.56	2.511
10121	eng	Brent	CHI	1;01.21	female	Child	61	68	1.115	0.367
010129	eng	Brent	OTH	.	.	Unidentified	13	29	2.231	1.25
10129	eng	Brent	MOT	.	female	Mother	840	3105	3.696	2.878
10129	eng	Brent	CHI	1;01.29	female	Child	53	64	1.208	0.406
010217	eng	Brent	MOT	.	female	Mother	990	3629	3.666	2.598
10217	eng	Brent	CHI	1;02.17	female	Child	81	100	1.235	0.452
010304	eng	Brent	OTH	.	.	Unidentified	4	15	3.75	1.479
10304	eng	Brent	MOT	.	female	Mother	674	2637	3.912	2.744
10304	eng	Brent	CHI	1;03.04	female	Child	40	48	1.2	0.458
average CHI									1.2556	0.6612
average MOT									4.032	2.81281818

A.2 Sample search results for "kind of" in Brent corpus

To decide whether a result in the corpus was generic, episodic, or ambiguous, the search output was saved as a .txt-file. Any result was referred to in the corpus (via the direct output in CLAN) and when in doubt, the audio-file was played on the CHILDES database online. The .txt-file was then colour-coded in

- red for episodic reference
- green for generic reference
- blue for usage ambiguous between generic and episodic reference.

The following is a sample of this kind of search for "kind of" in the Brent corpus:

combo +skind^of *.cha
Mon Oct 22 11:09:32 2018
combo (12-Sep-2018) is conducting analyses on:
  ALL speaker tiers
****
From file <000827.cha>

    Strings matched 0 times

From file <000828.cha>

    Strings matched 0 times

From file <000830 2.cha>

    Strings matched 0 times

From file <000830.cha>

    Strings matched 0 times

From file <000900.cha>

    Strings matched 0 times

From file <000901.cha>

    Strings matched 0 times

From file <000902.cha>

    Strings matched 0 times

From file <000904.cha>

    Strings matched 0 times

From file <000907.cha>

    Strings matched 0 times

From file <000908.cha>

    Strings matched 0 times

From file <000910.cha>

%%% EPISODIC
*** File "000910.cha": line 1097.
*MOT: (1)kind (1)of a neat toy huh ? %snd:"000910"_1671427_1672937

Strings matched 1 times

From file <000912.cha>

%%% EPISODIC
*** File "000912.cha": line 3660.
*MOT: you know (...) you're getting (1)kind (1)of heavy here [=! voice] .
%snd:"000912"_3966257_3968233

*** File "000912.cha": line 3665.
*MOT: yes you are getting (1)kind (1)of heavy . %snd:"000912"_3970343_3972520

Strings matched 2 times

From file <000914 2.cha>

Strings matched 0 times

From file <000914.cha>

%%%POTENTIALLY
*** File "000914.cha": line 699.
*MOT: that's a (1)kind (1)of a mixed bite . %snd:"000914"_656581_657788

Strings matched 1 times

From file <000917.cha>

*** File "000917.cha": line 1016.
*MOT: they're (1)kind (1)o(f) big balls . %snd:"000917"_969000_970268

*** File "000917.cha": line 1581.
*MOT: o:h it's (1)kind (1)o(f) chilly in here . %snd:"000917"_1538286_1539843

*** File "000917.cha": line 4225.
*MOT: what (1)kind (1)o(f) cereal do you want tomorrow , Morgan &=noise ?
%snd:"000917"_4036276_4038876

*** File "000917.cha": line 4475.
*MOT: they're (1)kind (1)of (...) butternutie@n in color &=noise .
%snd:"000917"_4484956_4488493

Strings matched 4 times

From file <000920.cha>

Strings matched 0 times

From file <000922.cha>

Strings matched 0 times

From file <000923.cha>

Strings matched 0 times

From file <000925.cha>

Strings matched 0 times

From file <000927.cha>

Strings matched 0 times

From file <000928 2.cha>

Strings matched 0 times

From file <000928.cha>

Strings matched 0 times

From file <000930.cha>

*** File "000930.cha": line 223.
*MOT: I could really do without this (.) (1)kind (1)of (..) technical stuff .
%snd:"000930"_157267_160612

*** File "000930.cha": line 469.
*MOT: (be)cause he makes wee+wees sometimes (.) (be)cause he's that (1)kind
(1)of a boy . %snd:"000930"_560051_563170

*** File "000930.cha": line 1333.
*MOT: I don't know which way you stepped on it but it actually (1)kind (1)of
bent it . %snd:"000930"_1408747_1411617

Strings matched 3 times

From file <001003.cha>

Strings matched 0 times

From file <001005 2.cha>

Strings matched 0 times

From file <001005.cha>

*** File "001005.cha": line 595.

*MOT: that's (...) (1)kind (1)of a . %snd:"001005"_429456_430582

*** File "001005.cha": line 1097.
*MOT: you can just (1)kind (1)of rove around and eat it &=noise .
%snd:"001005"_868970_870722

Strings matched 2 times

From file <001010.cha>

Strings matched 0 times

From file <001012.cha>

Strings matched 0 times

From file <001013 2.cha>

Strings matched 0 times

From file <001013.cha>

*** File "001013.cha": line 991.
*MOT: (1)kind (1)o(f) slides away from you on the floor huh &=noise ?
%snd:"001013"_1395670_1397667

*** File "001013.cha": line 1310.
*MOT: alright I'm getting (1)kind (1)o(f) tired of this whining .
%snd:"001013"_1723854_1726159

*** File "001013.cha": line 1315.
*MOT: I'm getting (1)kind (1)of tired of this constant whining .
%snd:"001013"_1727348_1729838

Strings matched 3 times

From file <001014.cha>

*** File "001014.cha": line 1391.
*MOT: yeah that's (1)kind (1)of close . %snd:"001014"_2310507_2312119

Strings matched 1 times

From file <001016.cha>

Strings matched 0 times

From file <001020.cha>

Strings matched 0 times

B

Stimuli for Pseudoword learning experiments

Contents

B.1 English pseudoword learning studies

B.1.1 Participant Information Sheet and consent form

Participants who were recruited via Amazon's Mechanical Turk were redirected to ibex-farm. They were presented with the following pages before starting the training phase, asking them to respond to a short socio-linguistic questionnaire and provide consent. They were then shown the instructions that informed them about the set-up of the study.

If you have not done so already, please read the informed consent. In lieu of your signature, please be sure to check the consent box below.

Before we begin, the researchers would like to know a little bit about you. Please fill in the following part.

Basic Information:

What is your age?	
With which gender do you most identify?	◯ Man ◯ Woman ◯ Trans Man ◯ Trans Woman ◯ Gender Non-Conforming ◯ Prefer Not to Answer Not Listed
What is your native language:	◯English ◯ Non-English
Did you live in an English-speaking environment from birth until (at least) age 13?	◯Yes ◯No
Did both of your parents speak English to you at home?	◯Yes ◯No
Which state did you grow up in (two letter abbreviation)?	-- select an option --
☐ I consent to participating in this experiment.	
Please enter your Prolific ID	

When you're ready, please turn off any distractions for the duration of the experiment!

→ Click here to continue

Instructions:

Welcome!

You have arrived on a foreign island and are surrounded by many unfamiliar objects, creatures, and other things.

Fortunately, you have met some researchers who will introduce you to some of the things around you, but these researchers also need your help in characterizing several new things too.

First, they will tell you what they know about 60 things they have found on the island. The things they have found are unusual, so the researchers have made up a vocabulary to describe them.

Pay close attention to what they tell you as this will help you think about the characteristics of the new things that you and the researchers find.

You don't have to remember what specific things are like. Rather, try to get a general feeling of what the island may be like and what kinds of characteristics things can have.

After this, the researchers will take you to a different part of the island. There, you will be asked what you think about the characteristics of the new things that you encounter.

→ Click here to continue

B.1.2 Sample trials

Below is an abbreviated version of the BP vs DS *pseudoword* study. This shows three sample training trials, where the property 'larbial' is trained with BP subjects and the property 'voove' with DS subjects. This is followed by the break message and three sample testing trials, where two questions target PCs (paraphrased with 'is an aspect of') and one question targets SCs (paraphrased with 'just because most'). The final page shows the optional comment box before participants received their completion code.

progress

Here we have a tematid.

Tematids are a larbial thing.

1. Press 1 or click to continue when you are ready.

progress

Here we have a lorfol.

The lorfol is a voove thing.

1. Press 1 or click to continue when you are ready.

progress

Ah! Here is a rodrup.

Rodrups are a larbial thing.

1. Press 1 or click to continue when you are ready.

progress

Good job! That was the first part. We will now move on to a different region of the island that the researchers haven't explored yet. They hope that the characteristics they have given you so far will be useful in helping them answer some questions about the new things they find there. The researchers will ask you to answer some questions about the things they encounter based on what you have learned so far.

You may not always feel very certain about your answer, so try to go with your initial reaction to the question.

Press 1 or the space bar to continue.

progress

Here's a cleckie and it has a larbial.

Do you think that it is an aspect of being a cleckie that it has a larbial?

(Definitely No)	1	2	3	4	5	6	7	*(Definitely Yes)*

progress

We thought we might find a tradub and this one has a larbial.

Do you think that tradubs have a larbial just because most tradubs have a larbial?

(Definitely No) 1 2 3 4 5 6 7 *(Definitely Yes)*

progress

Here's a hilome and it has a larbial.

Do you think that it is an aspect of being a hilome that it has a larbial?

(Definitely No)	1	2	3	4	5	6	7	*(Definitely Yes)*

OPTIONAL

Thank you for your participation. Before we get to the end of the experiment, we'd like to ask you one more question. Is there anything you'd like to tell the researchers or anything else that you noticed during the experiment?

Thank you for your participation! Your completion URL for Prolific is on the next page.

→ Click here to continue

B.1.3 Full list of stimuli

The tables on the following pages include the stimuli for both the training and the testing phase for all the English pseudoword learning studies presented in this thesis.

Stimuli for pseudoword learning studies

Experiment 1a (and 1a-R): Bare plural vs. indefinite singular subjects

Training stimuli (1a and 1a-R)

Ah! Here is a bebe.	Bebes have a voove.
Ah! Here is a bebe.	A bebe has a voove.
Here we have a cav.	Cavs have a voove.
Here we have a cav.	A cav has a voove.
Over here is a habin.	Habins have a voove.
Over here is a habin.	A habin has a voove.
If you look over here, you will see a jeev.	Jeevs have a voove.
If you look over here, you will see a jeev.	A jeev has a voove.
This is a keff.	Keffs have a voove.
This is a keff.	A keff has a voove.
That is a kirk.	Kirks have a voove.
That is a kirk.	A kirk has a voove.
Ah! Here is a leebel.	Leebels have a voove.
Ah! Here is a leebel.	A leebel has a voove.
Here we have a mallar.	Mallars have a voove.
Here we have a mallar.	A mallar has a voove.
Over here is a mogot.	Mogots have a voove.
Over here is a mogot.	A mogot has a voove.
If you look over here, you will see a morath.	Moraths have a voove.
If you look over here, you will see a morath.	A morath has a voove.
This is a teef.	Teefs have a voove.
This is a teef.	A teef has a voove.
That is a toogit.	Toogits have a voove.
That is a toogit.	A toogit has a voove.
Ah! Here is a winoozle.	Winoozles have a voove.
Ah! Here is a winoozle.	A winoozle has a voove.
Here we have a wug.	Wugs have a voove.
Here we have a wug.	A wug has a voove.
Over here is a zav.	Zavs have a voove.
Over here is a zav.	A zav has a voove.
If you look over here, you will see a basdo.	Basdos have a larbial.
If you look over here, you will see a basdo.	A basdo has a larbial.
This is a dohi.	Dohis have a larbial.
This is a dohi.	A dohi has a larbial.

That is a goulbap.	Goulbaps have a larbial.
That is a goulbap.	A goulbap has a larbial.
Ah! Here is a jospro.	Jospros have a larbial.
Ah! Here is a jospro.	A jospro has a larbial.
Here we have a jubchall.	Jubchalls have a larbial.
Here we have a jubchall.	A jubchall has a larbial.
Over here is a lezup.	Lezups have a larbial.
Over here is a lezup.	A lezup has a larbial.
If you look over here, you will see a namstol.	Namstols have a larbial.
If you look over here, you will see a namstol.	A namstol has a larbial.
This is a nedrill.	Nedrills have a larbial.
This is a nedrill.	A nedrill has a larbial.
That is a pruvia.	Pruvias have a larbial.
That is a pruvia.	A pruvia has a larbial.
Ah! Here is a rodrup.	Rodrups have a larbial.
Ah! Here is a rodrup.	A rodrup has a larbial.
Here we have a shurgin.	Shurgins have a larbial.
Here we have a shurgin.	A shurgin has a larbial.
Over here is a sorson.	Sorsons have a larbial.
Over here is a sorson.	A sorson has a larbial.
If you look over here, you will see a telato.	Telatos have a larbial.
If you look over here, you will see a telato.	A telato has a larbial.
This is a vallum.	Vallums have a larbial.
This is a vallum.	A vallum has a larbial.
That is a wulder.	Wulders have a larbial.
That is a wulder.	A wulder has a larbial.

Testing stimuli (1a)

Oh look! That's a cleckie and it has a larbial.	Do you think that it is an aspect of being a cleckie that it has a larbial?
We thought we might find a moff And this one has a larbial.	Do you think that it is an aspect of being a moff that it has a larbial?
Here's a moff and it has a larbial.	Do you think that it is an aspect of being a moff that it has a larbial?
Oh look! That's a moff and it has a larbial.	Do you think that it is an aspect of being a moff that it has a larbial?
Ah, we were hoping to find a moff, and look! It has a larbial.	Do you think that it is an aspect of being a moff that it has a larbial?
Here's a cleckie and it has a pilovier.	Do you think that it is an aspect of being a cleckie that it has a pilovier?
Ah, we were hoping to find a cleckie, and look! It has a pilovier.	Do you think that it is an aspect of being a cleckie that it has a pilovier?
Excellent! We've found a cleckie, and it has a pilovier.	Do you think that it is an aspect of being a cleckie that it has a pilovier?
We thought we might find a cleckie And this one has a pilovier.	Do you think that it is an aspect of being a cleckie that it has a pilovier?
Here's a shump and it has a pilovier.	Do you think that it is an aspect of being a shump that it has a pilovier?
Here's a gwooch and it has a pulappi.	Do you think that it is an aspect of being a gwooch that it has a pulappi?
Oh look! That's a shump and it has a pulappi.	Do you think that it is an aspect of being a shump that it has a pulappi?
Ah, we were hoping to find a shump, and look! It has a pulappi.	Do you think that it is an aspect of being a shump that it has a pulappi?
Excellent! We've found a shump, and it has a pulappi.	Do you think that it is an aspect of being a shump that it has a pulappi?
We thought we might find a shump And this one has a pulappi.	Do you think that it is an aspect of being a shump that it has a pulappi?
Excellent! We've found a gwooch, and it has a swoquix.	Do you think that it is an aspect of being a gwooch that it has a swoquix?
We thought we might find a gwooch And this one has a swoquix.	Do you think that it is an aspect of being a gwooch that it has a swoquix?
Oh look! That's a gwooch and it has a swoquix.	Do you think that it is an aspect of being a gwooch that it has a swoquix?
Ah, we were hoping to find a gwooch, and look! It has a swoquix.	Do you think that it is an aspect of being a gwooch that it has a swoquix?
Oh look! That's a pimsette and it has a swoquix.	Do you think that it is an aspect of being a pimsette that it has a swoquix?
Excellent! We've found a moff, and it has a voove.	Do you think that it is an aspect of being a moff that it has a voove?
Excellent! We've found a pimsette, and it has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?
We thought we might find a pimsette And this one has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?

Here's a pimsette and it has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?
Ah, we were hoping to find a pimsette, and look! It has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?
Oh look! That's a frorf and it has a larbial.	Do you think that it is not an aspect of being a frorf that it has a larbial?
Ah, we were hoping to find a kooze, and look! It has a larbial.	Do you think that it is not an aspect of being a kooze that it has a larbial?
Excellent! We've found a kooze, and it has a larbial.	Do you think that it is not an aspect of being a kooze that it has a larbial?
We thought we might find a kooze And this one has a larbial.	Do you think that it is not an aspect of being a kooze that it has a larbial?
Here's a kooze and it has a larbial.	Do you think that it is not an aspect of being a kooze that it has a larbial?
Here's a frorf and it has a pilovier.	Do you think that it is not an aspect of being a frorf that it has a pilovier?
Ah, we were hoping to find a frorf, and look! It has a pilovier.	Do you think that it is not an aspect of being a frorf that it has a pilovier?
Excellent! We've found a frorf, and it has a pilovier.	Do you think that it is not an aspect of being a frorf that it has a pilovier?
We thought we might find a frorf And this one has a pilovier.	Do you think that it is not an aspect of being a frorf that it has a pilovier?
Ah, we were hoping to find a tratt, and look! It has a pilovier.	Do you think that it is not an aspect of being a tratt that it has a pilovier?
Oh look! That's a rettle and it has a pulappi.	Do you think that it is not an aspect of being a rettle that it has a pulappi?
Excellent! We've found a tratt, and it has a pulappi.	Do you think that it is not an aspect of being a tratt that it has a pulappi?
We thought we might find a tratt And this one has a pulappi.	Do you think that it is not an aspect of being a tratt that it has a pulappi?
Here's a tratt and it has a pulappi.	Do you think that it is not an aspect of being a tratt that it has a pulappi?
Oh look! That's a tratt and it has a pulappi.	Do you think that it is not an aspect of being a tratt that it has a pulappi?
Ah, we were hoping to find a hink, and look! It has a swoquix.	Do you think that it is not an aspect of being a hink that it has a swoquix?
We thought we might find a rettle And this one has a swoquix.	Do you think that it is not an aspect of being a rettle that it has a swoquix?
Here's a rettle and it has a swoquix.	Do you think that it is not an aspect of being a rettle that it has a swoquix?
Ah, we were hoping to find a rettle, and look! It has a swoquix.	Do you think that it is not an aspect of being a rettle that it has a swoquix?
Excellent! We've found a rettle, and it has a swoquix.	Do you think that it is not an aspect of being a rettle that it has a swoquix?
We thought we might find a hink And this one has a voove.	Do you think that it is not an aspect of being a hink that it has a voove?
Here's a hink and it has a voove.	Do you think that it is not an aspect of being a hink that it has a voove?

Oh look! That's a hink and it has a voove.	Do you think that it is not an aspect of being a hink that it has a voove?
Excellent! We've found a hink, and it has a voove.	Do you think that it is not an aspect of being a hink that it has a voove?
Oh look! That's a kooze and it has a voove.	Do you think that it is not an aspect of being a kooze that it has a voove?
Oh look! That's a neeg and it has a larbial.	Do you think that neegs have a larbial, but not in virtue of being a neeg?
Ah, we were hoping to find a neeg, and look! It has a larbial.	Do you think that neegs have a larbial, but not in virtue of being a neeg?
Excellent! We've found a neeg, and it has a larbial.	Do you think that neegs have a larbial, but not in virtue of being a neeg?
We thought we might find a neeg And this one has a larbial.	Do you think that neegs have a larbial, but not in virtue of being a neeg?
Ah, we were hoping to find a throod, and look! It has a larbial.	Do you think that throods have a larbial, but not in virtue of being a throod?
Oh look! That's a blick and it has a pilovier.	Do you think that blicks have a pilovier, but not in virtue of being a blick?
Oh look! That's a throod and it has a pilovier.	Do you think that throods have a pilovier, but not in virtue of being a throod?
Excellent! We've found a throod, and it has a pilovier.	Do you think that throods have a pilovier, but not in virtue of being a throod?
We thought we might find a throod And this one has a pilovier.	Do you think that throods have a pilovier, but not in virtue of being a throod?
Here's a throod and it has a pilovier.	Do you think that throods have a pilovier, but not in virtue of being a throod?
Ah, we were hoping to find a blick, and look! It has a pulappi.	Do you think that blicks have a pulappi, but not in virtue of being a blick?
Excellent! We've found a blick, and it has a pulappi.	Do you think that blicks have a pulappi, but not in virtue of being a blick?
We thought we might find a blick And this one has a pulappi.	Do you think that blicks have a pulappi, but not in virtue of being a blick?
Here's a blick and it has a pulappi.	Do you think that blicks have a pulappi, but not in virtue of being a blick?
Ah, we were hoping to find a joof, and look! It has a pulappi.	Do you think that joofs have a pulappi, but not in virtue of being a joof?
We thought we might find a cleeg And this one has a swoquix.	Do you think that cleegs have a swoquix, but not in virtue of being a cleeg?
Here's a joof and it has a swoquix.	Do you think that joofs have a swoquix, but not in virtue of being a joof?
Oh look! That's a joof and it has a swoquix.	Do you think that joofs have a swoquix, but not in virtue of being a joof?
Excellent! We've found a joof, and it has a swoquix.	Do you think that joofs have a swoquix, but not in virtue of being a joof?
We thought we might find a joof And this one has a swoquix.	Do you think that joofs have a swoquix, but not in virtue of being a joof?
Oh look! That's a cleeg and it has a voove.	Do you think that cleegs have a voove, but not in virtue of being a cleeg?

Ah, we were hoping to find a cleeg, and look! It has a voove.	Do you think that cleegs have a voove, but not in virtue of being a cleeg?
Excellent! We've found a cleeg, and it has a voove.	Do you think that cleegs have a voove, but not in virtue of being a cleeg?
Here's a cleeg and it has a voove.	Do you think that cleegs have a voove, but not in virtue of being a cleeg?
Here's a neeg and it has a voove.	Do you think that neegs have a voove, but not in virtue of being a neeg?
Here's a narga and it has a larbial.	Do you think that a narga, in vitrue of being a narga, has a larbial?
Oh look! That's a teckle and it has a larbial.	Do you think that a teckle, in vitrue of being a teckle, has a larbial?
Ah, we were hoping to find a teckle, and look! It has a larbial.	Do you think that a teckle, in vitrue of being a teckle, has a larbial?
Excellent! We've found a teckle, and it has a larbial.	Do you think that a teckle, in vitrue of being a teckle, has a larbial?
We thought we might find a teckle And this one has a larbial.	Do you think that a teckle, in vitrue of being a teckle, has a larbial?
We thought we might find a narga And this one has a pilovier.	Do you think that a narga, in vitrue of being a narga, has a pilovier?
Oh look! That's a narga and it has a pilovier.	Do you think that a narga, in vitrue of being a narga, has a pilovier?
Ah, we were hoping to find a narga, and look! It has a pilovier.	Do you think that a narga, in vitrue of being a narga, has a pilovier?
Excellent! We've found a narga, and it has a pilovier.	Do you think that a narga, in vitrue of being a narga, has a pilovier?
Here's a peetle and it has a pilovier.	Do you think that a peetle, in vitrue of being a peetle, has a pilovier?
Here's a crerpin and it has a pulappi.	Do you think that a crerpin, in vitrue of being a crerpin, has a pulappi?
Oh look! That's a peetle and it has a pulappi.	Do you think that a peetle, in vitrue of being a peetle, has a pulappi?
Ah, we were hoping to find a peetle, and look! It has a pulappi.	Do you think that a peetle, in vitrue of being a peetle, has a pulappi?
Excellent! We've found a peetle, and it has a pulappi.	Do you think that a peetle, in vitrue of being a peetle, has a pulappi?
We thought we might find a peetle And this one has a pulappi.	Do you think that a peetle, in vitrue of being a peetle, has a pulappi?
Excellent! We've found a crerpin, and it has a swoquix.	Do you think that a crerpin, in vitrue of being a crerpin, has a swoquix?
We thought we might find a crerpin And this one has a swoquix.	Do you think that a crerpin, in vitrue of being a crerpin, has a swoquix?
Oh look! That's a crerpin and it has a swoquix.	Do you think that a crerpin, in vitrue of being a crerpin, has a swoquix?
Ah, we were hoping to find a crerpin, and look! It has a swoquix.	Do you think that a crerpin, in vitrue of being a crerpin, has a swoquix?
Ah, we were hoping to find a yoot, and look! It has a swoquix.	Do you think that a yoot, in vitrue of being a yoot, has a swoquix?

Here's a teckle and it has a voove.	Do you think that a teckle, in vitrue of being a teckle, has a voove?
We thought we might find a yoot And this one has a voove.	Do you think that a yoot, in vitrue of being a yoot, has a voove?
Here's a yoot and it has a voove.	Do you think that a yoot, in vitrue of being a yoot, has a voove?
Oh look! That's a yoot and it has a voove.	Do you think that a yoot, in vitrue of being a yoot, has a voove?
Excellent! We've found a yoot, and it has a voove.	Do you think that a yoot, in vitrue of being a yoot, has a voove?
Here's a chezzle and it has a larbial.	Do you think that chezzles have a larbial just because most chezzles have a larbial?
Excellent! We've found a clonkle, and it has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
We thought we might find a clonkle And this one has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
Here's a clonkle and it has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
Oh look! That's a clonkle and it has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
We thought we might find a chezzle And this one has a pilovier.	Do you think that chezzles have a pilovier just because most chezzles have a pilovier?
Oh look! That's a chezzle and it has a pilovier.	Do you think that chezzles have a pilovier just because most chezzles have a pilovier?
Ah, we were hoping to find a chezzle, and look! It has a pilovier.	Do you think that chezzles have a pilovier just because most chezzles have a pilovier?
Excellent! We've found a chezzle, and it has a pilovier.	Do you think that chezzles have a pilovier just because most chezzles have a pilovier?
Oh look! That's a pellop and it has a pilovier.	Do you think that pellops have a pilovier just because most pellops have a pilovier?
Ah, we were hoping to find a pellop, and look! It has a pulappi.	Do you think that pellops have a pulappi just because most pellops have a pulappi?
Excellent! We've found a pellop, and it has a pulappi.	Do you think that pellops have a pulappi just because most pellops have a pulappi?
We thought we might find a pellop And this one has a pulappi.	Do you think that pellops have a pulappi just because most pellops have a pulappi?
Here's a pellop and it has a pulappi.	Do you think that pellops have a pulappi just because most pellops have a pulappi?
We thought we might find a thunket And this one has a pulappi.	Do you think that thunkets have a pulappi just because most thunkets have a pulappi?
Oh look! That's a stirp and it has a swoquix.	Do you think that stirps have a swoquix just because most stirps have a swoquix?
Ah, we were hoping to find a thunket, and look! It has a swoquix.	Do you think that thunkets have a swoquix just because most thunkets have a swoquix?
Excellent! We've found a thunket, and it has a swoquix.	Do you think that thunkets have a swoquix just because most thunkets have a swoquix?
Here's a thunket and it has a swoquix.	Do you think that thunkets have a swoquix just because most thunkets have a swoquix?

Oh look! That's a thunket and it has a swoquix.	Do you think that thunkets have a swoquix just because most thunkets have a swoquix?
Ah, we were hoping to find a clonkle, and look! It has a voove.	Do you think that clonkles have a voove just because most clonkles have a voove?
Excellent! We've found a stirp, and it has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
We thought we might find a stirp And this one has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
Here's a stirp and it has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
Ah, we were hoping to find a stirp, and look! It has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
We thought we might find a flollo And this one has a larbial.	Do you think that flollos just happen to have a larbial?
Ah, we were hoping to find a gugget, and look! It has a larbial.	Do you think that guggets just happen to have a larbial?
Excellent! We've found a gugget, and it has a larbial.	Do you think that guggets just happen to have a larbial?
We thought we might find a gugget And this one has a larbial.	Do you think that guggets just happen to have a larbial?
Here's a gugget and it has a larbial.	Do you think that guggets just happen to have a larbial?
Excellent! We've found a flollo, and it has a pilovier.	Do you think that flollos just happen to have a pilovier?
Here's a flollo and it has a pilovier.	Do you think that flollos just happen to have a pilovier?
Oh look! That's a flollo and it has a pilovier.	Do you think that flollos just happen to have a pilovier?
Ah, we were hoping to find a flollo, and look! It has a pilovier.	Do you think that flollos just happen to have a pilovier?
Oh look! That's a lirpee and it has a pilovier.	Do you think that lirpees just happen to have a pilovier?
Oh look! That's a frelum and it has a pulappi.	Do you think that frelums just happen to have a pulappi?
Ah, we were hoping to find a lirpee, and look! It has a pulappi.	Do you think that lirpees just happen to have a pulappi?
Excellent! We've found a lirpee, and it has a pulappi.	Do you think that lirpees just happen to have a pulappi?
We thought we might find a lirpee And this one has a pulappi.	Do you think that lirpees just happen to have a pulappi?
Here's a lirpee and it has a pulappi.	Do you think that lirpees just happen to have a pulappi?
Excellent! We've found a drork, and it has a swoquix.	Do you think that drorks just happen to have a swoquix?
We thought we might find a frelum And this one has a swoquix.	Do you think that frelums just happen to have a swoquix?
Here's a frelum and it has a swoquix.	Do you think that frelums just happen to have a swoquix?

Ah, we were hoping to find a frelum, and look! It has a swoquix.	Do you think that frelums just happen to have a swoquix?
Excellent! We've found a frelum, and it has a swoquix.	Do you think that frelums just happen to have a swoquix?
Here's a drork and it has a voove.	Do you think that drorks just happen to have a voove?
Oh look! That's a drork and it has a voove.	Do you think that drorks just happen to have a voove?
Ah, we were hoping to find a drork, and look! It has a voove.	Do you think that drorks just happen to have a voove?
We thought we might find a drork And this one has a voove.	Do you think that drorks just happen to have a voove?
Oh look! That's a gugget and it has a voove.	Do you think that guggets just happen to have a voove?
Here's a dozzle and it has a larbial.	Do you think that dozzles have a larbial, but not just because most ,dozzles have a larbial?
Excellent! We've found a gwemb, and it has a larbial.	Do you think that gwembs have a larbial, but not just because most ,gwembs have a larbial?
We thought we might find a gwemb And this one has a larbial.	Do you think that gwembs have a larbial, but not just because most ,gwembs have a larbial?
Here's a gwemb and it has a larbial.	Do you think that gwembs have a larbial, but not just because most ,gwembs have a larbial?
Oh look! That's a gwemb and it has a larbial.	Do you think that gwembs have a larbial, but not just because most ,gwembs have a larbial?
Ah, we were hoping to find a blust, and look! It has a pilovier.	Do you think that blusts have a pilovier, but not just because most ,blusts have a pilovier?
We thought we might find a dozzle And this one has a pilovier.	Do you think that dozzles have a pilovier, but not just because most ,dozzles have a pilovier?
Oh look! That's a dozzle and it has a pilovier.	Do you think that dozzles have a pilovier, but not just because most ,dozzles have a pilovier?
Ah, we were hoping to find a dozzle, and look! It has a pilovier.	Do you think that dozzles have a pilovier, but not just because most ,dozzles have a pilovier?
Excellent! We've found a dozzle, and it has a pilovier.	Do you think that dozzles have a pilovier, but not just because most ,dozzles have a pilovier?
Excellent! We've found a blust, and it has a pulappi.	Do you think that blusts have a pulappi, but not just because most ,blusts have a pulappi?
We thought we might find a blust And this one has a pulappi.	Do you think that blusts have a pulappi, but not just because most ,blusts have a pulappi?
Here's a blust and it has a pulappi.	Do you think that blusts have a pulappi, but not just because most ,blusts have a pulappi?
Oh look! That's a blust and it has a pulappi.	Do you think that blusts have a pulappi, but not just because most ,blusts have a pulappi?
Excellent! We've found a drost, and it has a pulappi.	Do you think that drosts have a pulappi, but not just because most ,drosts have a pulappi?
Oh look! That's a drost and it has a swoquix.	Do you think that drosts have a swoquix, but not just because most ,drosts have a swoquix?
Ah, we were hoping to find a drost, and look! It has a swoquix.	Do you think that drosts have a swoquix, but not just because most ,drosts have a swoquix?

We thought we might find a drost And this one has a swoquix.	Do you think that drosts have a swoquix, but not just because most ,drosts have a swoquix?
Here's a drost and it has a swoquix.	Do you think that drosts have a swoquix, but not just because most ,drosts have a swoquix?
Excellent! We've found a rorpin, and it has a swoquix.	Do you think that rorpins have a swoquix, but not just because most ,rorpins have a swoquix?
Ah, we were hoping to find a gwemb, and look! It has a voove.	Do you think that gwembs have a voove, but not just because most ,gwembs have a voove?
Here's a rorpin and it has a voove.	Do you think that rorpins have a voove, but not just because most ,rorpins have a voove?
Oh look! That's a rorpin and it has a voove.	Do you think that rorpins have a voove, but not just because most ,rorpins have a voove?
Ah, we were hoping to find a rorpin, and look! It has a voove.	Do you think that rorpins have a voove, but not just because most ,rorpins have a voove?
We thought we might find a rorpin And this one has a voove.	Do you think that rorpins have a voove, but not just because most ,rorpins have a voove?
Excellent! We've found a fezzoo, and it has a larbial.	Do you think that fezzoos don't just happen to have a larbial?
Excellent! We've found a slend, and it has a larbial.	Do you think that slends don't just happen to have a larbial?
We thought we might find a slend And this one has a larbial.	Do you think that slends don't just happen to have a larbial?
Here's a slend and it has a larbial.	Do you think that slends don't just happen to have a larbial?
Oh look! That's a slend and it has a larbial.	Do you think that slends don't just happen to have a larbial?
Ah, we were hoping to find a fezzoo, and look! It has a pilovier.	Do you think that fezzoos don't just happen to have a pilovier?
We thought we might find a fezzoo And this one has a pilovier.	Do you think that fezzoos don't just happen to have a pilovier?
Here's a fezzoo and it has a pilovier.	Do you think that fezzoos don't just happen to have a pilovier?
Oh look! That's a fezzoo and it has a pilovier.	Do you think that fezzoos don't just happen to have a pilovier?
We thought we might find a zurp And this one has a pilovier.	Do you think that zurps don't just happen to have a pilovier?
Oh look! That's a swodum and it has a pulappi.	Do you think that swodums don't just happen to have a pulappi?
Here's a zurp and it has a pulappi.	Do you think that zurps don't just happen to have a pulappi?
Oh look! That's a zurp and it has a pulappi.	Do you think that zurps don't just happen to have a pulappi?
Ah, we were hoping to find a zurp, and look! It has a pulappi.	Do you think that zurps don't just happen to have a pulappi?
Excellent! We've found a zurp, and it has a pulappi.	Do you think that zurps don't just happen to have a pulappi?
We thought we might find a swodum And this one has a swoquix.	Do you think that swodums don't just happen to have a swoquix?

Here's a swodum and it has a swoquix.	Do you think that swodums don't just happen to have a swoquix?
Ah, we were hoping to find a swodum, and look! It has a swoquix.	Do you think that swodums don't just happen to have a swoquix?
Excellent! We've found a swodum, and it has a swoquix.	Do you think that swodums don't just happen to have a swoquix?
We thought we might find a zuzzette And this one has a swoquix.	Do you think that zuzzettes don't just happen to have a swoquix?
Ah, we were hoping to find a slend, and look! It has a voove.	Do you think that slends don't just happen to have a voove?
Oh look! That's a zuzzette and it has a voove.	Do you think that zuzzettes don't just happen to have a voove?
Ah, we were hoping to find a zuzzette, and look! It has a voove.	Do you think that zuzzettes don't just happen to have a voove?
Excellent! We've found a zuzzette, and it has a voove.	Do you think that zuzzettes don't just happen to have a voove?
Here's a zuzzette and it has a voove.	Do you think that zuzzettes don't just happen to have a voove?

Testing stimuli (1a-R)

Excellent! We've found a pimsette, and it has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?
Excellent! We've found a pimsette, and it has a larbial.	Do you think that it is an aspect of being a pimsette that it has a larbial?
Excellent! We've found a pimsette, and it has a larbial.	Do you think that it is an aspect of being a pimsette that it has a larbial?
Excellent! We've found a pimsette, and it has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?
Oh look! That's a gwooch and it has a voove.	Do you think that it is an aspect of being a gwooch that it has a voove?
Oh look! That's a gwooch and it has a larbial.	Do you think that it is an aspect of being a gwooch that it has a larbial?
Oh look! That's a gwooch and it has a larbial.	Do you think that it is an aspect of being a gwooch that it has a larbial?
Oh look! That's a gwooch and it has a voove.	Do you think that it is an aspect of being a gwooch that it has a voove?
Here's a cleckie and it has a voove.	Do you think that it is an aspect of being a cleckie that it has a voove?
Here's a cleckie and it has a larbial.	Do you think that it is an aspect of being a cleckie that it has a larbial?
Here's a cleckie and it has a larbial.	Do you think that it is an aspect of being a cleckie that it has a larbial?
Here's a cleckie and it has a voove.	Do you think that it is an aspect of being a cleckie that it has a voove?
We thought we might find a moff and this one has a voove.	Do you think that it is an aspect of being a moff that it has a voove?
We thought we might find a moff and this one has a labrial.	Do you think that it is an aspect of being a moff that it has a larbial?
We thought we might find a moff and this one has a larbial.	Do you think that it is an aspect of being a moff that it has a larbial?
We thought we might find a moff and this one has a voove.	Do you think that it is an aspect of being a moff that it has a voove?
Excellent! We've found a shump, and it has a voove.	Do you think that it is an aspect of being a shump that it has a voove?
Excellent! We've found a shump, and it has a larbial.	Do you think that it is an aspect of being a shump that it has a larbial?
Excellent! We've found a shump, and it has a larbial.	Do you think that it is an aspect of being a shump that it has a larbial?
Excellent! We've found a shump, and it has a voove.	Do you think that it is an aspect of being a shump that it has a voove?
Oh look! That's a bax and it has a voove.	Do you think that it is an aspect of being a bax that it has a voove?
Oh look! That's a bax and it has a larbial.	Do you think that it is an aspect of being a bax that it has a larbial?
Oh look! That's a bax and it has a larbial.	Do you think that it is an aspect of being a bax that it has a larbial?
Oh look! That's a bax and it has a voove.	Do you think that it is an aspect of being a bax that it has a voove?

Here's a hilome and it has a voove.	Do you think that it is an aspect of being a hilome that it has a voove?
Here's a hilome and it has a larbial.	Do you think that it is an aspect of being a hilome that it has a larbial?
Here's a hilome and it has a larbial.	Do you think that it is an aspect of being a hilome that it has a larbial?
Here's a hilome and it has a voove.	Do you think that it is an aspect of being a hilome that it has a voove?
We thought we might find a ilkonk and this one has a voove.	Do you think that it is an aspect of being a ilkonk that it has a voove?
We thought we might find a ilkonk and this one has a labrial.	Do you think that it is an aspect of being a ilkonk that it has a larbial?
We thought we might find a ilkonk and this one has a larbial.	Do you think that it is an aspect of being a ilkonk that it has a larbial?
We thought we might find a ilkonk and this one has a voove.	Do you think that it is an aspect of being a ilkonk that it has a voove?
Excellent! We've found a jepel, and it has a voove.	Do you think that it is an aspect of being a jepel that it has a voove?
Excellent! We've found a jepel, and it has a larbial.	Do you think that it is an aspect of being a jepel that it has a larbial?
Excellent! We've found a jepel, and it has a larbial.	Do you think that it is an aspect of being a jepel that it has a larbial?
Excellent! We've found a jepel, and it has a voove.	Do you think that it is an aspect of being a jepel that it has a voove?
Oh look! That's a bormtu and it has a voove.	Do you think that it is an aspect of being a bormtu that it has a voove?
Oh look! That's a bormtu and it has a larbial.	Do you think that it is an aspect of being a bormtu that it has a larbial?
Oh look! That's a bormtu and it has a larbial.	Do you think that it is an aspect of being a bormtu that it has a larbial?
Oh look! That's a bormtu and it has a voove.	Do you think that it is an aspect of being a bormtu that it has a voove?
Here's a ruftaz and it has a voove.	Do you think that it is an aspect of being a ruftaz that it has a voove?
Here's a ruftaz and it has a larbial.	Do you think that it is an aspect of being a ruftaz that it has a larbial?
Here's a ruftaz and it has a larbial.	Do you think that it is an aspect of being a ruftaz that it has a larbial?
Here's a ruftaz and it has a voove.	Do you think that it is an aspect of being a ruftaz that it has a voove?
We thought we might find a grusst and this one has a voove.	Do you think that it is an aspect of being a grusst that it has a voove?
We thought we might find a grusst and this one has a labrial.	Do you think that it is an aspect of being a grusst that it has a larbial?
We thought we might find a grusst and this one has a larbial.	Do you think that it is an aspect of being a grusst that it has a larbial?
We thought we might find a grusst and this one has a voove.	Do you think that it is an aspect of being a grusst that it has a voove?

Excellent! We've found a yoot, and it has a voove.	Do you think that a yoot, by virtue of being a yoot, has a voove?
Excellent! We've found a yoot, and it has a larbial.	Do you think that a yoot, by virtue of being a yoot, has a larbial?
Excellent! We've found a yoot, and it has a larbial.	Do you think that a yoot, by virtue of being a yoot, has a larbial?
Excellent! We've found a yoot, and it has a voove.	Do you think that a yoot, by virtue of being a yoot, has a voove?
Oh look! That's a peetle and it has a voove.	Do you think that a peetle, by virtue of being a peetle, has a voove?
Oh look! That's a peetle and it has a larbial.	Do you think that a peetle, by virtue of being a peetle, has a larbial?
Oh look! That's a peetle and it has a larbial.	Do you think that a peetle, by virtue of being a peetle, has a larbial?
Oh look! That's a peetle and it has a voove.	Do you think that a peetle, by virtue of being a peetle, has a voove?
Here's a crerpin and it has a voove.	Do you think that a crerpin, by virtue of being a crerpin, has a voove?
Here's a crerpin and it has a larbial.	Do you think that a crerpin, by virtue of being a crerpin, has a larbial?
Here's a crerpin and it has a larbial.	Do you think that a crerpin, by virtue of being a crerpin, has a larbial?
Here's a crerpin and it has a voove.	Do you think that a crerpin, by virtue of being a crerpin, has a voove?
We thought we might find a rooflar and this one has a voove.	Do you think that a rooflar, by virtue of being a rooflar, has a voove?
We thought we might find a rooflar and this one has a labrial.	Do you think that a rooflar, by virtue of being a rooflar, has a larbial?
We thought we might find a rooflar and this one has a larbial.	Do you think that a rooflar, by virtue of being a rooflar, has a larbial?
We thought we might find a rooflar and this one has a voove.	Do you think that a rooflar, by virtue of being a rooflar, has a voove?
Excellent! We've found a keefle, and it has a voove.	Do you think that a keefle, by virtue of being a keefle, has a voove?
Excellent! We've found a keefle, and it has a larbial.	Do you think that a keefle, by virtue of being a keefle, has a larbial?
Excellent! We've found a keefle, and it has a larbial.	Do you think that a keefle, by virtue of being a keefle, has a larbial?
Excellent! We've found a keefle, and it has a voove.	Do you think that a keefle, by virtue of being a keefle, has a voove?
Oh look! That's a narga and it has a voove.	Do you think that a narga, by virtue of being a narga, has a voove?
Oh look! That's a narga and it has a larbial.	Do you think that a narga, by virtue of being a narga, has a larbial?
Oh look! That's a narga and it has a larbial.	Do you think that a narga, by virtue of being a narga, has a larbial?
Oh look! That's a narga and it has a voove.	Do you think that a narga, by virtue of being a narga, has a voove?

Here's a teckle and it has a voove.	Do you think that a teckle, by virtue of being a teckle, has a voove?
Here's a teckle and it has a larbial.	Do you think that a teckle, by virtue of being a teckle, has a larbial?
Here's a teckle and it has a larbial.	Do you think that a teckle, by virtue of being a teckle, has a larbial?
Here's a teckle and it has a voove.	Do you think that a teckle, by virtue of being a teckle, has a voove?
We thought we might find a jallik and this one has a voove.	Do you think that a jallik, by virtue of being a jallik, has a voove?
We thought we might find a jallik and this one has a labrial.	Do you think that a jallik, by virtue of being a jallik, has a larbial?
We thought we might find a jallik and this one has a larbial.	Do you think that a jallik, by virtue of being a jallik, has a larbial?
We thought we might find a jallik and this one has a voove.	Do you think that a jallik, by virtue of being a jallik, has a voove?
Excellent! We've found a fihlam, and it has a voove.	Do you think that a fihlam, by virtue of being a fihlam, has a voove?
Excellent! We've found a fihlam, and it has a larbial.	Do you think that a fihlam, by virtue of being a fihlam, has a larbial?
Excellent! We've found a fihlam, and it has a larbial.	Do you think that a fihlam, by virtue of being a fihlam, has a larbial?
Excellent! We've found a fihlam, and it has a voove.	Do you think that a fihlam, by virtue of being a fihlam, has a voove?
Oh look! That's a hastip and it has a voove.	Do you think that a hastip, by virtue of being a hastip, has a voove?
Oh look! That's a hastip and it has a larbial.	Do you think that a hastip, by virtue of being a hastip, has a larbial?
Oh look! That's a hastip and it has a larbial.	Do you think that a hastip, by virtue of being a hastip, has a larbial?
Oh look! That's a hastip and it has a voove.	Do you think that a hastip, by virtue of being a hastip, has a voove?
Here's a polk and it has a voove.	Do you think that a polk, by virtue of being a polk, has a voove?
Here's a polk and it has a larbial.	Do you think that a polk, by virtue of being a polk, has a larbial?
Here's a polk and it has a larbial.	Do you think that a polk, by virtue of being a polk, has a larbial?
Here's a polk and it has a voove.	Do you think that a polk, by virtue of being a polk, has a voove?
We thought we might find a gouwt and this one has a voove.	Do you think that a gouwt, by virtue of being a gouwt, has a voove?
We thought we might find a gouwt and this one has a labrial.	Do you think that a gouwt, by virtue of being a gouwt, has a larbial?
We thought we might find a gouwt and this one has a larbial.	Do you think that a gouwt, by virtue of being a gouwt, has a larbial?
We thought we might find a gouwt and this one has a voove.	Do you think that a gouwt, by virtue of being a gouwt, has a voove?

Excellent! We've found a thunket, and it has a voove.	Do you think that thunkets have a voove just because most thunkets have a voove?
Excellent! We've found a thunket, and it has a larbial.	Do you think that thunkets have a larbial just because most thunkets have a larbial?
Excellent! We've found a thunket, and it has a larbial.	Do you think that thunkets have a larbial just because most thunkets have a larbial?
Excellent! We've found a thunket, and it has a voove.	Do you think that thunkets have a voove just because most thunkets have a voove?
Oh look! That's a valdin and it has a voove.	Do you think that valdins have a voove just because most valdins have a voove?
Oh look! That's a valdin and it has a larbial.	Do you think that valdins have a larbial just because most valdins have a larbial?
Oh look! That's a valdin and it has a larbial.	Do you think that valdins have a larbial just because most valdins have a larbial?
Oh look! That's a valdin and it has a voove.	Do you think that valdins have a voove just because most valdins have a voove?
Here's a stirp and it has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
Here's a stirp and it has a larbial.	Do you think that stirps have a larbial just because most stirps have a larbial?
Here's a stirp and it has a larbial.	Do you think that stirps have a larbial just because most stirps have a larbial?
Here's a stirp and it has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
We thought we might find a chezzle and this one has a voove.	Do you think that chezzles have a voove just because most chezzles have a voove?
We thought we might find a chezzle and this one has a labrial.	Do you think that chezzles have a larbial just because most chezzles have a larbial?
We thought we might find a chezzle and this one has a larbial.	Do you think that chezzles have a larbial just because most chezzles have a larbial?
We thought we might find a chezzle and this one has a voove.	Do you think that chezzles have a voove just because most chezzles have a voove?
Excellent! We've found a pellop, and it has a voove.	Do you think that pellops have a voove just because most pellops have a voove?
Excellent! We've found a pellop, and it has a larbial.	Do you think that pellops have a larbial just because most pellops have a larbial?
Excellent! We've found a pellop, and it has a larbial.	Do you think that pellops have a larbial just because most pellops have a larbial?
Excellent! We've found a pellop, and it has a voove.	Do you think that pellops have a voove just because most pellops have a voove?
Oh look! That's a restin and it has a voove.	Do you think that restins have a voove just because most restins have a voove?
Oh look! That's a restin and it has a larbial.	Do you think that restins have a larbial just because most restins have a larbial?
Oh look! That's a restin and it has a larbial.	Do you think that restins have a larbial just because most restins have a larbial?
Oh look! That's a restin and it has a voove.	Do you think that restins have a voove just because most restins have a voove?

Here's a clonkle and it has a voove.	Do you think that clonkles have a voove just because most clonkles have a voove?
Here's a clonkle and it has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
Here's a clonkle and it has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
Here's a clonkle and it has a voove.	Do you think that clonkles have a voove just because most clonkles have a voove?
We thought we might find a dreep and this one has a voove.	Do you think that dreeps have a voove just because most dreeps have a voove?
We thought we might find a dreep and this one has a labrial.	Do you think that dreeps have a larbial just because most dreeps have a larbial?
We thought we might find a dreep and this one has a larbial.	Do you think that dreeps have a larbial just because most dreeps have a larbial?
We thought we might find a dreep and this one has a voove.	Do you think that dreeps have a voove just because most dreeps have a voove?
Excellent! We've found a palpar, and it has a voove.	Do you think that palpars have a voove just because most palpars have a voove?
Excellent! We've found a palpar, and it has a larbial.	Do you think that palpars have a larbial just because most palpars have a larbial?
Excellent! We've found a palpar, and it has a larbial.	Do you think that palpars have a larbial just because most palpars have a larbial?
Excellent! We've found a palpar, and it has a voove.	Do you think that palpars have a voove just because most palpars have a voove?
Oh look! That's a baneak and it has a voove.	Do you think that baneaks have a voove just because most baneaks have a voove?
Oh look! That's a baneak and it has a larbial.	Do you think that baneaks have a larbial just because most baneaks have a larbial?
Oh look! That's a baneak and it has a larbial.	Do you think that baneaks have a larbial just because most baneaks have a larbial?
Oh look! That's a baneak and it has a voove.	Do you think that baneaks have a voove just because most baneaks have a voove?
Here's a yurtov and it has a voove.	Do you think that yurtovs have a voove just because most yurtovs have a voove?
Here's a yurtov and it has a larbial.	Do you think that yurtovs have a larbial just because most yurtovs have a larbial?
Here's a yurtov and it has a larbial.	Do you think that yurtovs have a larbial just because most yurtovs have a larbial?
Here's a yurtov and it has a voove.	Do you think that yurtovs have a voove just because most yurtovs have a voove?
We thought we might find a tradub and this one has a voove.	Do you think that tradubs have a voove just because most tradubs have a voove?
We thought we might find a tradub and this one has a labrial.	Do you think that tradubs have a larbial just because most tradubs have a larbial?
We thought we might find a tradub and this one has a larbial.	Do you think that tradubs have a larbial just because most tradubs have a larbial?
We thought we might find a tradub and this one has a voove.	Do you think that tradubs have a voove just because most tradubs have a voove?

Excellent! We've found a ziam, and it has a voove.	Do you think that ziams just happen to have a voove?
Excellent! We've found a ziam, and it has a larbial.	Do you think that ziams just happen to have a larbial?
Excellent! We've found a ziam, and it has a larbial.	Do you think that ziams just happen to have a larbial?
Excellent! We've found a ziam, and it has a voove.	Do you think that ziams just happen to have a voove?
Oh look! That's a nampa and it has a voove.	Do you think that nampas just happen to have a voove?
Oh look! That's a nampa and it has a larbial.	Do you think that nampas just happen to have a larbial?
Oh look! That's a nampa and it has a larbial.	Do you think that nampas just happen to have a larbial?
Oh look! That's a nampa and it has a voove.	Do you think that nampas just happen to have a voove?
Here's a drork and it has a voove.	Do you think that drorks just happen to have a voove?
Here's a drork and it has a larbial.	Do you think that drorks just happen to have a larbial?
Here's a drork and it has a larbial.	Do you think that drorks just happen to have a larbial?
Here's a drork and it has a voove.	Do you think that drorks just happen to have a voove?
We thought we might find a frelum and this one has a voove.	Do you think that frelums just happen to have a voove?
We thought we might find a frelum and this one has a labrial.	Do you think that frelums just happen to have a larbial?
We thought we might find a frelum and this one has a larbial.	Do you think that frelums just happen to have a larbial?
We thought we might find a frelum and this one has a voove.	Do you think that frelums just happen to have a voove?
Excellent! We've found a lirpee, and it has a voove.	Do you think that lirpees just happen to have a voove?
Excellent! We've found a lirpee, and it has a larbial.	Do you think that lirpees just happen to have a larbial?
Excellent! We've found a lirpee, and it has a larbial.	Do you think that lirpees just happen to have a larbial?
Excellent! We've found a lirpee, and it has a voove.	Do you think that lirpees just happen to have a voove?
Oh look! That's a flollo and it has a voove.	Do you think that flollos just happen to have a voove?
Oh look! That's a flollo and it has a larbial.	Do you think that flollos just happen to have a larbial?
Oh look! That's a flollo and it has a larbial.	Do you think that flollos just happen to have a larbial?
Oh look! That's a flollo and it has a voove.	Do you think that flollos just happen to have a voove?

Here's a gugget and it has a voove.	Do you think that guggets just happen to have a voove?
Here's a gugget and it has a larbial.	Do you think that guggets just happen to have a larbial?
Here's a gugget and it has a larbial.	Do you think that guggets just happen to have a larbial?
Here's a gugget and it has a voove.	Do you think that guggets just happen to have a voove?
We thought we might find a wullmon and this one has a voove.	Do you think that wullmons just happen to have a voove?
We thought we might find a wullmon and this one has a labrial.	Do you think that wullmons just happen to have a larbial?
We thought we might find a wullmon and this one has a larbial.	Do you think that wullmons just happen to have a larbial?
We thought we might find a wullmon and this one has a voove.	Do you think that wullmons just happen to have a voove?
Excellent! We've found a lolix, and it has a voove.	Do you think that lolixs just happen to have a voove?
Excellent! We've found a lolix, and it has a larbial.	Do you think that lolixs just happen to have a larbial?
Excellent! We've found a lolix, and it has a larbial.	Do you think that lolixs just happen to have a larbial?
Excellent! We've found a lolix, and it has a voove.	Do you think that lolixs just happen to have a voove?
Oh look! That's a gufnin and it has a voove.	Do you think that gufnins just happen to have a voove?
Oh look! That's a gufnin and it has a larbial.	Do you think that gufnins just happen to have a larbial?
Oh look! That's a gufnin and it has a larbial.	Do you think that gufnins just happen to have a larbial?
Oh look! That's a gufnin and it has a voove.	Do you think that gufnins just happen to have a voove?
Here's a rasdid and it has a voove.	Do you think that rasdids just happen to have a voove?
Here's a rasdid and it has a larbial.	Do you think that rasdids just happen to have a larbial?
Here's a rasdid and it has a larbial.	Do you think that rasdids just happen to have a larbial?
Here's a rasdid and it has a voove.	Do you think that rasdids just happen to have a voove?
We thought we might find a plifgir and this one has a voove.	Do you think that plifgirs just happen to have a voove?
We thought we might find a plifgir and this one has a labrial.	Do you think that plifgirs just happen to have a larbial?
We thought we might find a plifgir and this one has a larbial.	Do you think that plifgirs just happen to have a larbial?
We thought we might find a plifgir and this one has a voove.	Do you think that plifgirs just happen to have a voove?

Experiment 1b (and 1b-R): Bare plural vs. definite singular subjects

Training stimuli (1b and 1b-R)

Ah! Here is a bebe.	Bebes are a voove thing.
This is a bebe.	The bebe is a voove thing.
Here we have a cav.	Cavs are a voove thing.
Over here is a cav.	The cav is a voove thing.
Here we have a habin.	Habins are a voove thing.
Over here is a habin.	The habin is a voove thing.
If you look over here, you will see a jeev.	Jeevs are a voove thing.
Here we have a jeev.	The jeev is a voove thing.
This is a keff.	Keffs are a voove thing.
If you look over here, you will see a keff.	The keff is a voove thing.
That thing is a kirk.	Kirks are a voove thing.
Ah! Here is a kirk.	The kirk is a voove thing.
Here we have a leebel.	Leebels are a voove thing.
If you look over here, you will see a leebel.	The leebel is a voove thing.
Over here is a mallar.	Mallars are a voove thing.
That thing is a mallar.	The mallar is a voove thing.
That thing is a mogot.	Mogots are a voove thing.
Ah! Here is a mogot.	The mogot is a voove thing.
That thing is a morath.	Moraths are a voove thing.
Here we have a morath.	The morath is a voove thing.
Ah! Here is a teef.	Teefs are a voove thing.
This is a teef.	The teef is a voove thing.
If you look over here, you will see a toogit.	Toogits are a voove thing.
Ah! Here is a toogit.	The toogit is a voove thing.
Over here is a winoozle.	Winoozles are a voove thing.
Ah! Here is a winoozle.	The winoozle is a voove thing.
If you look over here, you will see a wug.	Wugs are a voove thing.
This is a wug.	The wug is a voove thing.
Over here is a zav.	Zavs are a voove thing.
That thing is a zav.	The zav is a voove thing.
This is a basdo.	Basdos are a larbial thing.
That thing is a basdo.	The basdo is a larbial thing.
This is a dohi.	Dohis are a larbial thing.
Ah! Here is a dohi.	The dohi is a larbial thing.
If you look over here, you will see a goulbap.	Goulbaps are a larbial thing.
This is a goulbap.	The goulbap is a larbial thing.

Ah! Here is a jospro.	Jospros are a larbial thing.
That thing is a jospro.	The jospro is a larbial thing.
If you look over here, you will see a jubchall.	Jubchalls are a larbial thing.
This is a jubchall.	The jubchall is a larbial thing.
Here we have a lezup.	Lezups are a larbial thing.
If you look over here, you will see a lezup.	The lezup is a larbial thing.
That thing is a namstol.	Namstols are a larbial thing.
Over here is a namstol.	The namstol is a larbial thing.
This is a nedrill.	Nedrills are a larbial thing.
Here we have a nedrill.	The nedrill is a larbial thing.
If you look over here, you will see a pruvia.	Pruvias are a larbial thing.
Here we have a pruvia.	The pruvia is a larbial thing.
This is a rodrup.	Rodrups are a larbial thing.
Ah! Here is a rodrup.	The rodrup is a larbial thing.
That thing is a shurgin.	Shurgins are a larbial thing.
Over here is a shurgin.	The shurgin is a larbial thing.
Here we have a sorson.	Sorsons are a larbial thing.
Over here is a sorson.	The sorson is a larbial thing.
Ah! Here is a telato.	Telatos are a larbial thing.
That thing is a telato.	The telato is a larbial thing.
If you look over here, you will see a vallum.	Vallums are a larbial thing.
Over here is a vallum.	The vallum is a larbial thing.
Over here is a wulder.	Wulders are a larbial thing.
Here we have a wulder.	The wulder is a larbial thing.

Testing stimuli (1b)

Excellent! We've found a pimsette, and it is voove.	Do you think that it is an aspect of being a pimsette that it is voove?
Here's a cleckie and it is voove.	Do you think that it is an aspect of being a cleckie that it is voove?
Oh look! That's a gwooch and it is larbial.	Do you think that it is an aspect of being a gwooch that it is larbial?
Here's a shump and it is larbial.	Do you think that it is an aspect of being a shump that it is larbial?
Oh look! That's a hink and it is voove.	Do you think that it is not an aspect of being a hink that it is voove?
Ah, we were hoping to find a frorf, and look! It is voove.	Do you think that it is not an aspect of being a frorf that it is voove?
Ah, we were hoping to find a rettle, and look! It is larbial.	Do you think that it is not an aspect of being a rettle that it is larbial?
Oh look! That's a tratt and it is larbial.	Do you think that it is not an aspect of being a tratt that it is larbial?
Here's a cleeg and it is voove.	Do you think that cleegs are voove, but not in virtue of being a cleeg?
We thought we might find a joof. And this one is voove.	Do you think that joofs are voove, but not in virtue of being a joof?
Ah, we were hoping to find a blick, and look! It is larbial.	Do you think that blicks are larbial, but not in virtue of being a blick?
Ah, we were hoping to find a neeg, and look! It is larbial.	Do you think that neegs are larbial, but not in virtue of being a neeg?
Here's a yoot and it is voove.	Do you think that a yoot, in vitrue of being a yoot, is voove?
Ah, we were hoping to find a crerpin, and look! It is voove.	Do you think that a crerpin, in vitrue of being a crerpin, is voove?
Here's a peetle and it is larbial.	Do you think that a peetle, in vitrue of being a peetle, is larbial?
Here's a teckle and it is larbial.	Do you think that a teckle, in vitrue of being a teckle, is larbial?
Ah, we were hoping to find a stirp, and look! It is voove.	Do you think that stirps are voove just because most stirps are voove?
Oh look! That's a pellop and it is voove.	Do you think that pellops are voove just because most pellops are voove?
Oh look! That's a thunket and it is larbial.	Do you think that thunkets are larbial just because most thunkets are larbial?
We thought we might find a pellop. And this one is larbial.	Do you think that pellops are larbial just because most pellops are larbial?
Oh look! That's a flollo and it is voove.	Do you think that flollos just happen to be voove?
Excellent! We've found a frelum, and it is voove.	Do you think that frelums just happen to be voove?
Oh look! That's a lirpee and it is larbial.	Do you think that lirpees just happen to be larbial?
We thought we might find a gugget. And this one is larbial.	Do you think that guggets just happen to be larbial?

Ah, we were hoping to find a rorpin, and look! It is voove.	Do you think that rorpins are voove, but not just because most rorpins are voove?
Ah, we were hoping to find a dozzle, and look! It is voove.	Do you think that dozzles are voove, but not just because most dozzles are voove?
Excellent! We've found a drost, and it is larbial.	Do you think that drosts are larbial, but not just because most drosts are larbial?
Excellent! We've found a blust, and it is larbial.	Do you think that blusts are larbial, but not just because most blusts are larbial?
Ah, we were hoping to find a zuzzette, and look! It is voove.	Do you think that zuzzettes don't just happen to be voove?
We thought we might find a swodum. And this one is voove.	Do you think that swodums don't just happen to be voove?
Oh look! That's a zurp and it is larbial.	Do you think that zurps don't just happen to be larbial?
Excellent! We've found a slend, and it is larbial.	Do you think that slends don't just happen to be larbial?
We thought we might find a yoot And this one has a voove.	Do you think that a yoot, in vitrue of being a yoot, has a voove?
Excellent! We've found a pimsette, and it has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?
Oh look! That's a cleeg and it has a voove.	Do you think that cleegs have a voove, but not in virtue of being a cleeg?
We thought we might find a hink And this one has a voove.	Do you think that it is not an aspect of being a hink that it has a voove?
Here's a drork and it has a voove.	Do you think that drorks just happen to have a voove?
Excellent! We've found a stirp, and it has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
Oh look! That's a zuzzette and it has a voove.	Do you think that zuzzettes don't just happen to have a voove?
Here's a rorpin and it has a voove.	Do you think that rorpins have a voove, but not just because most ,rorpins have a voove?
Excellent! We've found a narga, and it has a pilovier.	Do you think that a narga, in vitrue of being a narga, has a pilovier?
We thought we might find a cleckie And this one has a pilovier.	Do you think that it is an aspect of being a cleckie that it has a pilovier?
Here's a throod and it has a pilovier.	Do you think that throods have a pilovier, but not in virtue of being a throod?
We thought we might find a frorf And this one has a pilovier.	Do you think that it is not an aspect of being a frorf that it has a pilovier?
Ah, we were hoping to find a flollo, and look! It has a pilovier.	Do you think that flollos just happen to have a pilovier?
Excellent! We've found a chezzle, and it has a pilovier.	Do you think that chezzles have a pilovier just because most chezzles have a pilovier?
Oh look! That's a fezzoo and it has a pilovier.	Do you think that fezzoos don't just happen to have a pilovier?
Excellent! We've found a dozzle, and it has a pilovier.	Do you think that dozzles have a pilovier, but not just because most ,dozzles have a pilovier?

Oh look! That's a crerpin and it has a swoquix.	Do you think that a crerpin, in vitrue of being a crerpin, has a swoquix?
Oh look! That's a gwooch and it has a swoquix.	Do you think that it is an aspect of being a gwooch that it has a swoquix?
Excellent! We've found a joof, and it has a swoquix.	Do you think that joofs have a swoquix, but not in virtue of being a joof?
Ah, we were hoping to find a rettle, and look! It has a swoquix.	Do you think that it is not an aspect of being a rettle that it has a swoquix?
Ah, we were hoping to find a frelum, and look! It has a swoquix.	Do you think that frelums just happen to have a swoquix?
Here's a thunket and it has a swoquix.	Do you think that thunkets have a swoquix just because most thunkets have a swoquix?
Ah, we were hoping to find a swodum, and look! It has a swoquix.	Do you think that swodums don't just happen to have a swoquix?
We thought we might find a drost And this one has a swoquix.	Do you think that drosts have a swoquix, but not just because most ,drosts have a swoquix?
Excellent! We've found a peetle, and it has a pulappi.	Do you think that a peetle, in vitrue of being a peetle, has a pulappi?
Excellent! We've found a shump, and it has a pulappi.	Do you think that it is an aspect of being a shump that it has a pulappi?
We thought we might find a blick And this one has a pulappi.	Do you think that blicks have a pulappi, but not in virtue of being a blick?
Here's a tratt and it has a pulappi.	Do you think that it is not an aspect of being a tratt that it has a pulappi?
We thought we might find a lirpee And this one has a pulappi.	Do you think that lirpees just happen to have a pulappi?
We thought we might find a pellop And this one has a pulappi.	Do you think that pellops have a pulappi just because most pellops have a pulappi?
Ah, we were hoping to find a zurp, and look! It has a pulappi.	Do you think that zurps don't just happen to have a pulappi?
Here's a blust and it has a pulappi.	Do you think that blusts have a pulappi, but not just because most ,blusts have a pulappi?
Oh look! That's a teckle and it has a larbial.	Do you think that a teckle, in vitrue of being a teckle, has a larbial?
We thought we might find a moff And this one has a larbial.	Do you think that it is an aspect of being a moff that it has a larbial?
Oh look! That's a neeg and it has a larbial.	Do you think that neegs have a larbial, but not in virtue of being a neeg?
Ah, we were hoping to find a kooze, and look! It has a larbial.	Do you think that it is not an aspect of being a kooze that it has a larbial?
Ah, we were hoping to find a gugget, and look! It has a larbial.	Do you think that guggets just happen to have a larbial?
Excellent! We've found a clonkle, and it has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
Excellent! We've found a slend, and it has a larbial.	Do you think that slends don't just happen to have a larbial?
Excellent! We've found a gwemb, and it has a larbial.	Do you think that gwembs have a larbial, but not just because most ,gwembs have a larbial?

Here's a yoot and it has a voove.	Do you think that a yoot, in vitrue of being a yoot, has a voove?
We thought we might find a pimsette And this one has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?
Ah, we were hoping to find a cleeg, and look! It has a voove.	Do you think that cleegs have a voove, but not in virtue of being a cleeg?
Here's a hink and it has a voove.	Do you think that it is not an aspect of being a hink that it has a voove?
Oh look! That's a drork and it has a voove.	Do you think that drorks just happen to have a voove?
We thought we might find a stirp And this one has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
Ah, we were hoping to find a zuzzette, and look! It has a voove.	Do you think that zuzzettes don't just happen to have a voove?
Oh look! That's a rorpin and it has a voove.	Do you think that rorpins have a voove, but not just because most ,rorpins have a voove?
Oh look! That's a narga and it has a pilovier.	Do you think that a narga, in vitrue of being a narga, has a pilovier?
Ah, we were hoping to find a cleckie, and look! It has a pilovier.	Do you think that it is an aspect of being a cleckie that it has a pilovier?
Excellent! We've found a throod, and it has a pilovier.	Do you think that throods have a pilovier, but not in virtue of being a throod?
Ah, we were hoping to find a frorf, and look! It has a pilovier.	Do you think that it is not an aspect of being a frorf that it has a pilovier?
Here's a flollo and it has a pilovier.	Do you think that flollos just happen to have a pilovier?
Oh look! That's a chezzle and it has a pilovier.	Do you think that chezzles have a pilovier just because most chezzles have a pilovier?
We thought we might find a fezzoo And this one has a pilovier.	Do you think that fezzoos don't just happen to have a pilovier?
Oh look! That's a dozzle and it has a pilovier.	Do you think that dozzles have a pilovier, but not just because most ,dozzles have a pilovier?
Excellent! We've found a crerpin, and it has a swoquix.	Do you think that a crerpin, in vitrue of being a crerpin, has a swoquix?
Excellent! We've found a gwooch, and it has a swoquix.	Do you think that it is an aspect of being a gwooch that it has a swoquix?
Here's a joof and it has a swoquix.	Do you think that joofs have a swoquix, but not in virtue of being a joof?
We thought we might find a rettle And this one has a swoquix.	Do you think that it is not an aspect of being a rettle that it has a swoquix?
We thought we might find a frelum And this one has a swoquix.	Do you think that frelums just happen to have a swoquix?
Ah, we were hoping to find a thunket, and look! It has a swoquix.	Do you think that thunkets have a swoquix just because most thunkets have a swoquix?
We thought we might find a swodum And this one has a swoquix.	Do you think that swodums don't just happen to have a swoquix?

Oh look! That's a drost and it has a swoquix.	Do you think that drosts have a swoquix, but not just because most ,drosts have a swoquix?
We thought we might find a peetle And this one has a pulappi.	Do you think that a peetle, in vitrue of being a peetle, has a pulappi?
We thought we might find a shump And this one has a pulappi.	Do you think that it is an aspect of being a shump that it has a pulappi?
Here's a blick and it has a pulappi.	Do you think that blicks have a pulappi, but not in virtue of being a blick?
Oh look! That's a tratt and it has a pulappi.	Do you think that it is not an aspect of being a tratt that it has a pulappi?
Here's a lirpee and it has a pulappi.	Do you think that lirpees just happen to have a pulappi?
Here's a pellop and it has a pulappi.	Do you think that pellops have a pulappi just because most pellops have a pulappi?
Excellent! We've found a zurp, and it has a pulappi.	Do you think that zurps don't just happen to have a pulappi?
Oh look! That's a blust and it has a pulappi.	Do you think that blusts have a pulappi, but not just because most ,blusts have a pulappi?
We thought we might find a teckle And this one has a larbial.	Do you think that a teckle, in vitrue of being a teckle, has a larbial?
Ah, we were hoping to find a moff, and look! It has a larbial.	Do you think that it is an aspect of being a moff that it has a larbial?
We thought we might find a neeg And this one has a larbial.	Do you think that neegs have a larbial, but not in virtue of being a neeg?
Here's a kooze and it has a larbial.	Do you think that it is not an aspect of being a kooze that it has a larbial?
Here's a gugget and it has a larbial.	Do you think that guggets just happen to have a larbial?
Oh look! That's a clonkle and it has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
Oh look! That's a slend and it has a larbial.	Do you think that slends don't just happen to have a larbial?
Oh look! That's a gwemb and it has a larbial.	Do you think that gwembs have a larbial, but not just because most ,gwembs have a larbial?
Excellent! We've found a yoot, and it has a voove.	Do you think that a yoot, in vitrue of being a yoot, has a voove?
Ah, we were hoping to find a pimsette, and look! It has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?
Here's a cleeg and it has a voove.	Do you think that cleegs have a voove, but not in virtue of being a cleeg?
Excellent! We've found a hink, and it has a voove.	Do you think that it is not an aspect of being a hink that it has a voove?
We thought we might find a drork And this one has a voove.	Do you think that drorks just happen to have a voove?
Ah, we were hoping to find a stirp, and look! It has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
Here's a zuzzette and it has a voove.	Do you think that zuzzettes don't just happen to have a voove?

We thought we might find a rorpin And this one has a voove.	Do you think that rorpins have a voove, but not just because most ,rorpins have a voove?
Ah, we were hoping to find a narga, and look! It has a pilovier.	Do you think that a narga, in vitrue of being a narga, has a pilovier?
Excellent! We've found a cleckie, and it has a pilovier.	Do you think that it is an aspect of being a cleckie that it has a pilovier?
We thought we might find a throod And this one has a pilovier.	Do you think that throods have a pilovier, but not in virtue of being a throod?
Excellent! We've found a frorf, and it has a pilovier.	Do you think that it is not an aspect of being a frorf that it has a pilovier?
Oh look! That's a flollo and it has a pilovier.	Do you think that flollos just happen to have a pilovier?
Ah, we were hoping to find a chezzle, and look! It has a pilovier.	Do you think that chezzles have a pilovier just because most chezzles have a pilovier?
Here's a fezzoo and it has a pilovier.	Do you think that fezzoos don't just happen to have a pilovier?
Ah, we were hoping to find a dozzle, and look! It has a pilovier.	Do you think that dozzles have a pilovier, but not just because most ,dozzles have a pilovier?
We thought we might find a crerpin And this one has a swoquix.	Do you think that a crerpin, in vitrue of being a crerpin, has a swoquix?
We thought we might find a gwooch And this one has a swoquix.	Do you think that it is an aspect of being a gwooch that it has a swoquix?
Oh look! That's a joof and it has a swoquix.	Do you think that joofs have a swoquix, but not in virtue of being a joof?
Here's a rettle and it has a swoquix.	Do you think that it is not an aspect of being a rettle that it has a swoquix?
Here's a frelum and it has a swoquix.	Do you think that frelums just happen to have a swoquix?
Excellent! We've found a thunket, and it has a swoquix.	Do you think that thunkets have a swoquix just because most thunkets have a swoquix?
Here's a swodum and it has a swoquix.	Do you think that swodums don't just happen to have a swoquix?
Ah, we were hoping to find a drost, and look! It has a swoquix.	Do you think that drosts have a swoquix, but not just because most ,drosts have a swoquix?
Oh look! That's a peetle and it has a pulappi.	Do you think that a peetle, in vitrue of being a peetle, has a pulappi?
Oh look! That's a shump and it has a pulappi.	Do you think that it is an aspect of being a shump that it has a pulappi?
Ah, we were hoping to find a blick, and look! It has a pulappi.	Do you think that blicks have a pulappi, but not in virtue of being a blick?
Excellent! We've found a tratt, and it has a pulappi.	Do you think that it is not an aspect of being a tratt that it has a pulappi?
Ah, we were hoping to find a lirpee, and look! It has a pulappi.	Do you think that lirpees just happen to have a pulappi?
Ah, we were hoping to find a pellop, and look! It has a pulappi.	Do you think that pellops have a pulappi just because most pellops have a pulappi?

Here's a zurp and it has a pulappi.	Do you think that zurps don't just happen to have a pulappi?
Excellent! We've found a blust, and it has a pulappi.	Do you think that blusts have a pulappi, but not just because most ,blusts have a pulappi?
Ah, we were hoping to find a teckle, and look! It has a larbial.	Do you think that a teckle, in vitrue of being a teckle, has a larbial?
Here's a moff and it has a larbial.	Do you think that it is an aspect of being a moff that it has a larbial?
Ah, we were hoping to find a neeg, and look! It has a larbial.	Do you think that neegs have a larbial, but not in virtue of being a neeg?
Excellent! We've found a kooze, and it has a larbial.	Do you think that it is not an aspect of being a kooze that it has a larbial?
Excellent! We've found a gugget, and it has a larbial.	Do you think that guggets just happen to have a larbial?
We thought we might find a clonkle And this one has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
We thought we might find a slend And this one has a larbial.	Do you think that slends don't just happen to have a larbial?
We thought we might find a gwemb And this one has a larbial.	Do you think that gwembs have a larbial, but not just because most ,gwembs have a larbial?
Oh look! That's a yoot and it has a voove.	Do you think that a yoot, in vitrue of being a yoot, has a voove?
Here's a pimsette and it has a voove.	Do you think that it is an aspect of being a pimsette that it has a voove?
Excellent! We've found a cleeg, and it has a voove.	Do you think that cleegs have a voove, but not in virtue of being a cleeg?
Oh look! That's a hink and it has a voove.	Do you think that it is not an aspect of being a hink that it has a voove?
Ah, we were hoping to find a drork, and look! It has a voove.	Do you think that drorks just happen to have a voove?
Here's a stirp and it has a voove.	Do you think that stirps have a voove just because most stirps have a voove?
Excellent! We've found a zuzzette, and it has a voove.	Do you think that zuzzettes don't just happen to have a voove?
Ah, we were hoping to find a rorpin, and look! It has a voove.	Do you think that rorpins have a voove, but not just because most ,rorpins have a voove?
We thought we might find a narga And this one has a pilovier.	Do you think that a narga, in vitrue of being a narga, has a pilovier?
Here's a cleckie and it has a pilovier.	Do you think that it is an aspect of being a cleckie that it has a pilovier?
Oh look! That's a throod and it has a pilovier.	Do you think that throods have a pilovier, but not in virtue of being a throod?
Here's a frorf and it has a pilovier.	Do you think that it is not an aspect of being a frorf that it has a pilovier?
Excellent! We've found a flollo, and it has a pilovier.	Do you think that flollos just happen to have a pilovier?
We thought we might find a chezzle And this one has a pilovier.	Do you think that chezzles have a pilovier just because most chezzles have a pilovier?

Ah, we were hoping to find a fezzoo, and look! It has a pilovier.	Do you think that fezzoos don't just happen to have a pilovier?
We thought we might find a dozzle And this one has a pilovier.	Do you think that dozzles have a pilovier, but not just because most ,dozzles have a pilovier?
Ah, we were hoping to find a crerpin, and look! It has a swoquix.	Do you think that a crerpin, in vitrue of being a crerpin, has a swoquix?
Ah, we were hoping to find a gwooch, and look! It has a swoquix.	Do you think that it is an aspect of being a gwooch that it has a swoquix?
We thought we might find a joof And this one has a swoquix.	Do you think that joofs have a swoquix, but not in virtue of being a joof?
Excellent! We've found a rettle, and it has a swoquix.	Do you think that it is not an aspect of being a rettle that it has a swoquix?
Excellent! We've found a frelum, and it has a swoquix.	Do you think that frelums just happen to have a swoquix?
Oh look! That's a thunket and it has a swoquix.	Do you think that thunkets have a swoquix just because most thunkets have a swoquix?
Excellent! We've found a swodum, and it has a swoquix.	Do you think that swodums don't just happen to have a swoquix?
Here's a drost and it has a swoquix.	Do you think that drosts have a swoquix, but not just because most ,drosts have a swoquix?
Ah, we were hoping to find a peetle, and look! It has a pulappi.	Do you think that a peetle, in vitrue of being a peetle, has a pulappi?
Ah, we were hoping to find a shump, and look! It has a pulappi.	Do you think that it is an aspect of being a shump that it has a pulappi?
Excellent! We've found a blick, and it has a pulappi.	Do you think that blicks have a pulappi, but not in virtue of being a blick?
We thought we might find a tratt And this one has a pulappi.	Do you think that it is not an aspect of being a tratt that it has a pulappi?
Excellent! We've found a lirpee, and it has a pulappi.	Do you think that lirpees just happen to have a pulappi?
Excellent! We've found a pellop, and it has a pulappi.	Do you think that pellops have a pulappi just because most pellops have a pulappi?
Oh look! That's a zurp and it has a pulappi.	Do you think that zurps don't just happen to have a pulappi?
We thought we might find a blust And this one has a pulappi.	Do you think that blusts have a pulappi, but not just because most ,blusts have a pulappi?
Excellent! We've found a teckle, and it has a larbial.	Do you think that a teckle, in vitrue of being a teckle, has a larbial?
Oh look! That's a moff and it has a larbial.	Do you think that it is an aspect of being a moff that it has a larbial?
Excellent! We've found a neeg, and it has a larbial.	Do you think that neegs have a larbial, but not in virtue of being a neeg?
We thought we might find a kooze And this one has a larbial.	Do you think that it is not an aspect of being a kooze that it has a larbial?
We thought we might find a gugget And this one has a larbial.	Do you think that guggets just happen to have a larbial?

Here's a clonkle and it has a larbial.	Do you think that clonkles have a larbial just because most clonkles have a larbial?
Here's a slend and it has a larbial.	Do you think that slends don't just happen to have a larbial?
Here's a gwemb and it has a larbial.	Do you think that gwembs have a larbial, but not just because most ,gwembs have a larbial?

Testing stimuli (1b-R)

Excellent! We've found a pimsette, and it is a voove thing.	Do you think that it is an aspect of being a pimsette that it is a voove thing?
Excellent! We've found a pimsette, and it is a larbial thing.	Do you think that it is an aspect of being a pimsette that it is a larbial thing?
Excellent! We've found a pimsette, and it is a larbial thing.	Do you think that it is an aspect of being a pimsette that it is a larbial thing?
Excellent! We've found a pimsette, and it is a voove thing.	Do you think that it is an aspect of being a pimsette that it is a voove thing?
Here's a gwooch and it is a voove thing.	Do you think that it is an aspect of being a gwooch that it is a voove thing?
Here's a gwooch and it is a larbial thing.	Do you think that it is an aspect of being a gwooch that it is a larbial thing?
Here's a gwooch and it is a larbial thing.	Do you think that it is an aspect of being a gwooch that it is a larbial thing?
Here's a gwooch and it is a voove thing.	Do you think that it is an aspect of being a gwooch that it is a voove thing?
Oh look! That's a cleckie and it is a voove thing.	Do you think that it is an aspect of being a cleckie that it is a voove thing?
Oh look! That's a cleckie and it is a larbial thing.	Do you think that it is an aspect of being a cleckie that it is a larbial thing?
Oh look! That's a cleckie and it is a larbial thing.	Do you think that it is an aspect of being a cleckie that it is a larbial thing?
Oh look! That's a cleckie and it is a voove thing.	Do you think that it is an aspect of being a cleckie that it is a voove thing?
Ah, we were hoping to find a moff, and look! It is a voove thing.	Do you think that it is an aspect of being a moff that it is a voove thing?
Ah, we were hoping to find a moff, and look! It is a larbial thing.	Do you think that it is an aspect of being a moff that it is a larbial thing?
Ah, we were hoping to find a moff, and look! It is a larbial thing.	Do you think that it is an aspect of being a moff that it is a larbial thing?
Ah, we were hoping to find a moff, and look! It is a voove thing.	Do you think that it is an aspect of being a moff that it is a voove thing?
Excellent! We've found a shump, and it is a voove thing.	Do you think that it is an aspect of being a shump that it is a voove thing?
Excellent! We've found a shump, and it is a larbial thing.	Do you think that it is an aspect of being a shump that it is a larbial thing?
Excellent! We've found a shump, and it is a larbial thing.	Do you think that it is an aspect of being a shump that it is a larbial thing?
Excellent! We've found a shump, and it is a voove thing.	Do you think that it is an aspect of being a shump that it is a voove thing?
Here's a bax and it is a voove thing.	Do you think that it is an aspect of being a bax that it is a voove thing?
Here's a bax and it is a larbial thing.	Do you think that it is an aspect of being a bax that it is a larbial thing?
Here's a bax and it is a larbial thing.	Do you think that it is an aspect of being a bax that it is a larbial thing?
Here's a bax and it is a voove thing.	Do you think that it is an aspect of being a bax that it is a voove thing?

Oh look! That's a hilome and it is a voove thing.	Do you think that it is an aspect of being a hilome that it is a voove thing?
Oh look! That's a hilome and it is a larbial thing.	Do you think that it is an aspect of being a hilome that it is a larbial thing?
Oh look! That's a hilome and it is a larbial thing.	Do you think that it is an aspect of being a hilome that it is a larbial thing?
Oh look! That's a hilome and it is a voove thing.	Do you think that it is an aspect of being a hilome that it is a voove thing?
Ah, we were hoping to find a ilkonk, and look! It is a voove thing.	Do you think that it is an aspect of being a ilkonk that it is a voove thing?
Ah, we were hoping to find a ilkonk, and look! It is a larbial thing.	Do you think that it is an aspect of being a ilkonk that it is a larbial thing?
Ah, we were hoping to find a ilkonk, and look! It is a larbial thing.	Do you think that it is an aspect of being a ilkonk that it is a larbial thing?
Ah, we were hoping to find a ilkonk, and look! It is a voove thing.	Do you think that it is an aspect of being a ilkonk that it is a voove thing?
Excellent! We've found a jepel, and it is a voove thing.	Do you think that it is an aspect of being a jepel that it is a voove thing?
Excellent! We've found a jepel, and it is a larbial thing.	Do you think that it is an aspect of being a jepel that it is a larbial thing?
Excellent! We've found a jepel, and it is a larbial thing.	Do you think that it is an aspect of being a jepel that it is a larbial thing?
Excellent! We've found a jepel, and it is a voove thing.	Do you think that it is an aspect of being a jepel that it is a voove thing?
Here's a bormtu and it is a voove thing.	Do you think that it is an aspect of being a bormtu that it is a voove thing?
Here's a bormtu and it is a larbial thing.	Do you think that it is an aspect of being a bormtu that it is a larbial thing?
Here's a bormtu and it is a larbial thing.	Do you think that it is an aspect of being a bormtu that it is a larbial thing?
Here's a bormtu and it is a voove thing.	Do you think that it is an aspect of being a bormtu that it is a voove thing?
Oh look! That's a ruftaz and it is a voove thing.	Do you think that it is an aspect of being a ruftaz that it is a voove thing?
Oh look! That's a ruftaz and it is a larbial thing.	Do you think that it is an aspect of being a ruftaz that it is a larbial thing?
Oh look! That's a ruftaz and it is a larbial thing.	Do you think that it is an aspect of being a ruftaz that it is a larbial thing?
Oh look! That's a ruftaz and it is a voove thing.	Do you think that it is an aspect of being a ruftaz that it is a voove thing?
Ah, we were hoping to find a grusst, and look! It is a voove thing.	Do you think that it is an aspect of being a grusst that it is a voove thing?
Ah, we were hoping to find a grusst, and look! It is a larbial thing.	Do you think that it is an aspect of being a grusst that it is a larbial thing?
Ah, we were hoping to find a grusst, and look! It is a larbial thing.	Do you think that it is an aspect of being a grusst that it is a larbial thing?
Ah, we were hoping to find a grusst, and look! It is a voove thing.	Do you think that it is an aspect of being a grusst that it is a voove thing?

Excellent! We've found a yoot, and it is a voove thing.	Do you think that a yoot, by virtue of being a yoot, is a voove thing?
Excellent! We've found a yoot, and it is a larbial thing.	Do you think that a yoot, by virtue of being a yoot, is a larbial thing?
Excellent! We've found a yoot, and it is a larbial thing.	Do you think that a yoot, by virtue of being a yoot, is a larbial thing?
Excellent! We've found a yoot, and it is a voove thing.	Do you think that a yoot, by virtue of being a yoot, is a voove thing?
Here's a peetle and it is a voove thing.	Do you think that a peetle, by virtue of being a peetle, is a voove thing?
Here's a peetle and it is a larbial thing.	Do you think that a peetle, by virtue of being a peetle, is a larbial thing?
Here's a peetle and it is a larbial thing.	Do you think that a peetle, by virtue of being a peetle, is a larbial thing?
Here's a peetle and it is a voove thing.	Do you think that a peetle, by virtue of being a peetle, is a voove thing?
Oh look! That's a crerpin and it is a voove thing.	Do you think that a crerpin, by virtue of being a crerpin, is a voove thing?
Oh look! That's a crerpin and it is a larbial thing.	Do you think that a crerpin, by virtue of being a crerpin, is a larbial thing?
Oh look! That's a crerpin and it is a larbial thing.	Do you think that a crerpin, by virtue of being a crerpin, is a larbial thing?
Oh look! That's a crerpin and it is a voove thing.	Do you think that a crerpin, by virtue of being a crerpin, is a voove thing?
Ah, we were hoping to find a rooflar, and look! It is a voove thing.	Do you think that a rooflar, by virtue of being a rooflar, is a voove thing?
Ah, we were hoping to find a rooflar, and look! It is a larbial thing.	Do you think that a rooflar, by virtue of being a rooflar, is a larbial thing?
Ah, we were hoping to find a rooflar, and look! It is a larbial thing.	Do you think that a rooflar, by virtue of being a rooflar, is a larbial thing?
Ah, we were hoping to find a rooflar, and look! It is a voove thing.	Do you think that a rooflar, by virtue of being a rooflar, is a voove thing?
Excellent! We've found a keefle, and it is a voove thing.	Do you think that a keefle, by virtue of being a keefle, is a voove thing?
Excellent! We've found a keefle, and it is a larbial thing.	Do you think that a keefle, by virtue of being a keefle, is a larbial thing?
Excellent! We've found a keefle, and it is a larbial thing.	Do you think that a keefle, by virtue of being a keefle, is a larbial thing?
Excellent! We've found a keefle, and it is a voove thing.	Do you think that a keefle, by virtue of being a keefle, is a voove thing?
Here's a narga and it is a voove thing.	Do you think that a narga, by virtue of being a narga, is a voove thing?
Here's a narga and it is a larbial thing.	Do you think that a narga, by virtue of being a narga, is a larbial thing?
Here's a narga and it is a larbial thing.	Do you think that a narga, by virtue of being a narga, is a larbial thing?
Here's a narga and it is a voove thing.	Do you think that a narga, by virtue of being a narga, is a voove thing?

Oh look! That's a teckle and it is a voove thing.	Do you think that a teckle, by virtue of being a teckle, is a voove thing?
Oh look! That's a teckle and it is a larbial thing.	Do you think that a teckle, by virtue of being a teckle, is a larbial thing?
Oh look! That's a teckle and it is a larbial thing.	Do you think that a teckle, by virtue of being a teckle, is a larbial thing?
Oh look! That's a teckle and it is a voove thing.	Do you think that a teckle, by virtue of being a teckle, is a voove thing?
Ah, we were hoping to find a jallik, and look! It is a voove thing.	Do you think that a jallik, by virtue of being a jallik, is a voove thing?
Ah, we were hoping to find a jallik, and look! It is a larbial thing.	Do you think that a jallik, by virtue of being a jallik, is a larbial thing?
Ah, we were hoping to find a jallik, and look! It is a larbial thing.	Do you think that a jallik, by virtue of being a jallik, is a larbial thing?
Ah, we were hoping to find a jallik, and look! It is a voove thing.	Do you think that a jallik, by virtue of being a jallik, is a voove thing?
Excellent! We've found a fihlam, and it is a voove thing.	Do you think that a fihlam, by virtue of being a fihlam, is a voove thing?
Excellent! We've found a fihlam, and it is a larbial thing.	Do you think that a fihlam, by virtue of being a fihlam, is a larbial thing?
Excellent! We've found a fihlam, and it is a larbial thing.	Do you think that a fihlam, by virtue of being a fihlam, is a larbial thing?
Excellent! We've found a fihlam, and it is a voove thing.	Do you think that a fihlam, by virtue of being a fihlam, is a voove thing?
Here's a hastip and it is a voove thing.	Do you think that a hastip, by virtue of being a hastip, is a voove thing?
Here's a hastip and it is a larbial thing.	Do you think that a hastip, by virtue of being a hastip, is a larbial thing?
Here's a hastip and it is a larbial thing.	Do you think that a hastip, by virtue of being a hastip, is a larbial thing?
Here's a hastip and it is a voove thing.	Do you think that a hastip, by virtue of being a hastip, is a voove thing?
Oh look! That's a polk and it is a voove thing.	Do you think that a polk, by virtue of being a polk, is a voove thing?
Oh look! That's a polk and it is a larbial thing.	Do you think that a polk, by virtue of being a polk, is a larbial thing?
Oh look! That's a polk and it is a larbial thing.	Do you think that a polk, by virtue of being a polk, is a larbial thing?
Oh look! That's a polk and it is a voove thing.	Do you think that a polk, by virtue of being a polk, is a voove thing?
Ah, we were hoping to find a gouwt, and look! It is a voove thing.	Do you think that a gouwt, by virtue of being a gouwt, is a voove thing?
Ah, we were hoping to find a gouwt, and look! It is a larbial thing.	Do you think that a gouwt, by virtue of being a gouwt, is a larbial thing?
Ah, we were hoping to find a gouwt, and look! It is a larbial thing.	Do you think that a gouwt, by virtue of being a gouwt, is a larbial thing?
Ah, we were hoping to find a gouwt, and look! It is a voove thing.	Do you think that a gouwt, by virtue of being a gouwt, is a voove thing?

Excellent! We've found a thunket, and it is a voove thing.	Do you think that thunkets are a voove thing, just because most thunkets are a voove thing?
Excellent! We've found a thunket, and it is a larbial thing.	Do you think that thunkets are a larbial thing, just because most thunkets are a larbial thing?
Excellent! We've found a thunket, and it is a larbial thing.	Do you think that thunkets are a larbial thing, just because most thunkets are a larbial thing?
Excellent! We've found a thunket, and it is a voove thing.	Do you think that thunkets are a voove thing, just because most thunkets are a voove thing?
Here's a valdin and it is a voove thing.	Do you think that valdins are a voove thing, just because most valdins are a voove thing?
Here's a valdin and it is a larbial thing.	Do you think that valdins are a larbial thing, just because most valdins are a larbial thing?
Here's a valdin and it is a larbial thing.	Do you think that valdins are a larbial thing, just because most valdins are a larbial thing?
Here's a valdin and it is a voove thing.	Do you think that valdins are a voove thing, just because most valdins are a voove thing?
Oh look! That's a stirp and it is a voove thing.	Do you think that stirps are a voove thing, just because most stirps are a voove thing?
Oh look! That's a stirp and it is a larbial thing.	Do you think that stirps are a larbial thing, just because most stirps are a larbial thing?
Oh look! That's a stirp and it is a larbial thing.	Do you think that stirps are a larbial thing, just because most stirps are a larbial thing?
Oh look! That's a stirp and it is a voove thing.	Do you think that stirps are a voove thing, just because most stirps are a voove thing?
Ah, we were hoping to find a chezzle, and look! It is a voove thing.	Do you think that chezzles are a voove thing, just because most chezzles are a voove thing?
Ah, we were hoping to find a chezzle, and look! It is a larbial thing.	Do you think that chezzles are a larbial thing, just because most chezzles are a larbial thing?
Ah, we were hoping to find a chezzle, and look! It is a larbial thing.	Do you think that chezzles are a larbial thing, just because most chezzles are a larbial thing?
Ah, we were hoping to find a chezzle, and look! It is a voove thing.	Do you think that chezzles are a voove thing, just because most chezzles are a voove thing?
Excellent! We've found a pellop, and it is a voove thing.	Do you think that pellops are a voove thing, just because most pellops are a voove thing?
Excellent! We've found a pellop, and it is a larbial thing.	Do you think that pellops are a larbial thing, just because most pellops are a larbial thing?
Excellent! We've found a pellop, and it is a larbial thing.	Do you think that pellops are a larbial thing, just because most pellops are a larbial thing?
Excellent! We've found a pellop, and it is a voove thing.	Do you think that pellops are a voove thing, just because most pellops are a voove thing?
Here's a restin and it is a voove thing.	Do you think that restins are a voove thing, just because most restins are a voove thing?
Here's a restin and it is a larbial thing.	Do you think that restins are a larbial thing, just because most restins are a larbial thing?
Here's a restin and it is a larbial thing.	Do you think that restins are a larbial thing, just because most restins are a larbial thing?
Here's a restin and it is a voove thing.	Do you think that restins are a voove thing, just because most restins are a voove thing?

Oh look! That's a clonkle and it is a voove thing.	Do you think that clonkles are a voove thing, just because most clonkles are a voove thing?
Oh look! That's a clonkle and it is a larbial thing.	Do you think that clonkles are a larbial thing, just because most clonkles are a larbial thing?
Oh look! That's a clonkle and it is a larbial thing.	Do you think that clonkles are a larbial thing, just because most clonkles are a larbial thing?
Oh look! That's a clonkle and it is a voove thing.	Do you think that clonkles are a voove thing, just because most clonkles are a voove thing?
Ah, we were hoping to find a dreep, and look! It is a voove thing.	Do you think that dreeps are a voove thing, just because most dreeps are a voove thing?
Ah, we were hoping to find a dreep, and look! It is a larbial thing.	Do you think that dreeps are a larbial thing, just because most dreeps are a larbial thing?
Ah, we were hoping to find a dreep, and look! It is a larbial thing.	Do you think that dreeps are a larbial thing, just because most dreeps are a larbial thing?
Ah, we were hoping to find a dreep, and look! It is a voove thing.	Do you think that dreeps are a voove thing, just because most dreeps are a voove thing?
Excellent! We've found a palpar, and it is a voove thing.	Do you think that palpars are a voove thing, just because most palpars are a voove thing?
Excellent! We've found a palpar, and it is a larbial thing.	Do you think that palpars are a larbial thing, just because most palpars are a larbial thing?
Excellent! We've found a palpar, and it is a larbial thing.	Do you think that palpars are a larbial thing, just because most palpars are a larbial thing?
Excellent! We've found a palpar, and it is a voove thing.	Do you think that palpars are a voove thing, just because most palpars are a voove thing?
Here's a baneak and it is a voove thing.	Do you think that baneaks are a voove thing, just because most baneaks are a voove thing?
Here's a baneak and it is a larbial thing.	Do you think that baneaks are a larbial thing, just because most baneaks are a larbial thing?
Here's a baneak and it is a larbial thing.	Do you think that baneaks are a larbial thing, just because most baneaks are a larbial thing?
Here's a baneak and it is a voove thing.	Do you think that baneaks are a voove thing, just because most baneaks are a voove thing?
Oh look! That's a yurtov and it is a voove thing.	Do you think that yurtovs are a voove thing, just because most yurtovs are a voove thing?
Oh look! That's a yurtov and it is a larbial thing.	Do you think that yurtovs are a larbial thing, just because most yurtovs are a larbial thing?
Oh look! That's a yurtov and it is a larbial thing.	Do you think that yurtovs are a larbial thing, just because most yurtovs are a larbial thing?
Oh look! That's a yurtov and it is a voove thing.	Do you think that yurtovs are a voove thing, just because most yurtovs are a voove thing?
Ah, we were hoping to find a tradub, and look! It is a voove thing.	Do you think that tradubs are a voove thing, just because most tradubs are a voove thing?
Ah, we were hoping to find a tradub, and look! It is a larbial thing.	Do you think that tradubs are a larbial thing, just because most tradubs are a larbial thing?
Ah, we were hoping to find a tradub, and look! It is a larbial thing.	Do you think that tradubs are a larbial thing, just because most tradubs are a larbial thing?
Ah, we were hoping to find a tradub, and look! It is a voove thing.	Do you think that tradubs are a voove thing, just because most tradubs are a voove thing?

Excellent! We've found a ziam, and it is a voove thing.	Do you think that ziams just happen to be a voove thing?
Excellent! We've found a ziam, and it is a larbial thing.	Do you think that ziams just happen to be a larbial thing?
Excellent! We've found a ziam, and it is a larbial thing.	Do you think that ziams just happen to be a larbial thing?
Excellent! We've found a ziam, and it is a voove thing.	Do you think that ziams just happen to be a voove thing?
Here's a nampa and it is a voove thing.	Do you think that nampas just happen to be a voove thing?
Here's a nampa and it is a larbial thing.	Do you think that nampas just happen to be a larbial thing?
Here's a nampa and it is a larbial thing.	Do you think that nampas just happen to be a larbial thing?
Here's a nampa and it is a voove thing.	Do you think that nampas just happen to be a voove thing?
Oh look! That's a drork and it is a voove thing.	Do you think that drorks just happen to be a voove thing?
Oh look! That's a drork and it is a larbial thing.	Do you think that drorks just happen to be a larbial thing?
Oh look! That's a drork and it is a larbial thing.	Do you think that drorks just happen to be a larbial thing?
Oh look! That's a drork and it is a voove thing.	Do you think that drorks just happen to be a voove thing?
Ah, we were hoping to find a frelum, and look! It is a voove thing.	Do you think that frelums just happen to be a voove thing?
Ah, we were hoping to find a frelum, and look! It is a larbial thing.	Do you think that frelums just happen to be a larbial thing?
Ah, we were hoping to find a frelum, and look! It is a larbial thing.	Do you think that frelums just happen to be a larbial thing?
Ah, we were hoping to find a frelum, and look! It is a voove thing.	Do you think that frelums just happen to be a voove thing?
Excellent! We've found a lirpee, and it is a voove thing.	Do you think that lirpees just happen to be a voove thing?
Excellent! We've found a lirpee, and it is a larbial thing.	Do you think that lirpees just happen to be a larbial thing?
Excellent! We've found a lirpee, and it is a larbial thing.	Do you think that lirpees just happen to be a larbial thing?
Excellent! We've found a lirpee, and it is a voove thing.	Do you think that lirpees just happen to be a voove thing?
Here's a flollo and it is a voove thing.	Do you think that flollos just happen to be a voove thing?
Here's a flollo and it is a larbial thing.	Do you think that flollos just happen to be a larbial thing?
Here's a flollo and it is a larbial thing.	Do you think that flollos just happen to be a larbial thing?
Here's a flollo and it is a voove thing.	Do you think that flollos just happen to be a voove thing?

Oh look! That's a gugget and it is a voove thing.	Do you think that guggets just happen to be a voove thing?
Oh look! That's a gugget and it is a larbial thing.	Do you think that guggets just happen to be a larbial thing?
Oh look! That's a gugget and it is a larbial thing.	Do you think that guggets just happen to be a larbial thing?
Oh look! That's a gugget and it is a voove thing.	Do you think that guggets just happen to be a voove thing?
Ah, we were hoping to find a wullmon, and look! It is a voove thing.	Do you think that wullmons just happen to be a voove thing?
Ah, we were hoping to find a wullmon, and look! It is a larbial thing.	Do you think that wullmons just happen to be a larbial thing?
Ah, we were hoping to find a wullmon, and look! It is a larbial thing.	Do you think that wullmons just happen to be a larbial thing?
Ah, we were hoping to find a wullmon, and look! It is a voove thing.	Do you think that wullmons just happen to be a voove thing?
Excellent! We've found a lolix, and it is a voove thing.	Do you think that lolixs just happen to be a voove thing?
Excellent! We've found a lolix, and it is a larbial thing.	Do you think that lolixs just happen to be a larbial thing?
Excellent! We've found a lolix, and it is a larbial thing.	Do you think that lolixs just happen to be a larbial thing?
Excellent! We've found a lolix, and it is a voove thing.	Do you think that lolixs just happen to be a voove thing?
Here's a gufnin and it is a voove thing.	Do you think that gufnins just happen to be a voove thing?
Here's a gufnin and it is a larbial thing.	Do you think that gufnins just happen to be a larbial thing?
Here's a gufnin and it is a larbial thing.	Do you think that gufnins just happen to be a larbial thing?
Here's a gufnin and it is a voove thing.	Do you think that gufnins just happen to be a voove thing?
Oh look! That's a rasdid and it is a voove thing.	Do you think that rasdids just happen to be a voove thing?
Oh look! That's a rasdid and it is a larbial thing.	Do you think that rasdids just happen to be a larbial thing?
Oh look! That's a rasdid and it is a larbial thing.	Do you think that rasdids just happen to be a larbial thing?
Oh look! That's a rasdid and it is a voove thing.	Do you think that rasdids just happen to be a voove thing?
Ah, we were hoping to find a plifgir, and look! It is a voove thing.	Do you think that plifgirs just happen to be a voove thing?
Ah, we were hoping to find a plifgir, and look! It is a larbial thing.	Do you think that plifgirs just happen to be a larbial thing?

Ah, we were hoping to find a plifgir, and look! It is a larbial thing.	Do you think that plifgirs just happen to be a larbial thing?
Ah, we were hoping to find a plifgir, and look! It is a voove thing.	Do you think that plifgirs just happen to be a voove thing?

NB: The stimuli for Experiments 1c, 1d, and 1e are based on these training and testing stimuli and available on request. The stimuli for 1d and 1e include twice the amount of training sentences. I have included below a list of the pseudowords that were used in subject position.

Full list of pseudowords for Experiments 1d and 1e

bebe	basdo
cav	dohi
habin	goulbap
jeev	jospro
keff	jubchall
kirk	lezup
leebel	namstol
mallar	nedrill
mogot	pruvia
morath	rodrup
teef	shurgin
toogit	sorson
winoozle	telato
wug	vallum
zav	wulder
throod	tratt
rorpin	zuzet
drost	sutub
slend	grantim
blust	proccit
norp	suster
fauze	limprod
lorfol	norteb
nargan	holeelan
wirtun	vidulum
frikled	tematid
brifter	gritbel
gratig	rognop
vencon	nadinap
ricate	densuk

B.2 French pseudoword learning studies

B.2.1 Participant Information Sheet and consent form

LANGUAGE AND BRAIN LABORATORY

FACULTY OF LINGUISTICS, PHILOLOGY, AND PHONETICS

UNIVERSITY OF OXFORD

Dr E. Matthew Husband

matthew.husband@ling-phil.ox.ac.uk

37 Wellington Sq., Oxford, OX1 2JF

Tel: +44 (0)1865 280400 Fax: +44 (0)1865 280412

brainlab.clp.ox.ac.uk

Formulaire de consentement

TITRE DE L'ÉTUDE: Comment la compréhension du langage se reflète dans le comportement et le cerveau

DÉTAILS DU CHERCHEUR: Dr E. Matthew Husband, Associate Professor en Psycholinguistique dans la Faculté de Linguistique, Philologie et Phonétique

BUT DE L'ÉTUDE: Le but de ce projet de recherche est d'explorer expérimentalement différents modèles et théories de la compréhension du langage en ayant recours aux techniques de pointe

1. Je confirme que j'ai lu la fiche d'information et que j'ai eu l'occasion de poser toutes mes questions, auxquelles j'ai reçu des réponses satisfaisantes.
2. Je comprends que ce projet de recherche a été examiné par *The University of Oxford Central University Research Ethics Committee*, et en a obtenu l'approbation éthique.
3. Je comprends que ma participation à cette étude est volontaire et que je suis libre de m'en retirer et d'en retirer mes données personnelles en tout temps, sans indiquer de motif, et sans aucune conséquence négative ou pénalité.
4. Je comprends qui aura accès aux données à caractère personnel que je fournis dans le cadre de cette étude.
5. Je comprends comment mes données personnelles seront stockées, par exemple conformément à la Loi sur la Protection des Données, et ce qu'il en adviendra à la fin de cette étude.
6. Je comprends comment cette recherche sera rédigée et publiée dans des articles de journaux évalués par les pairs, dans des actes de colloque, dans les archives en ligne de l'Université ou dans des manuscrits à diffusion interne.
7. Je comprends comment transmettre une préoccupation ou déposer une plainte.
8. J'accepte de participer à cette étude.

En guise de signature, assurez-vous de bien cocher la case ci-dessous.

☐ Je consens à participer à cette expérience.

→ Click here to continue

Si vous ne l'avez pas déjà fait, veuillez lire le consentement éclairé.

Avant de commencer, les chercheurs aimeraient en savoir un peu plus sur vous. Veuillez renseigner les informations suivantes.

Informations de base::

Quel âge avez-vous?	
Quel est votre sexe biologique?	Homme Femme
Quelle est votre langue maternelle?	français non français
Avez-vous vécu dans un environnement français de la naissance à au moins 13 ans?	Oui Non
Est-ce que vos deux parents vous ont parlé français à la maison?	Oui Non
Dans quel pays avez-vous grandi?	

Lorsque vous êtes prêt, détournez-vous de toute distraction pendant toute la durée de l'expérience!

→ Click here to continue

B.2.2 Sample trials

Below is an abbreviated version of the IS vs DPl French *pseudoword* study. This shows the welcome instructions to participants, two sample training trials, where the property 'un vitin' is trained with IS subjects and the property 'une sépite' with DPl subjects. This is followed by the break message and four sample testing trials, where two questions target PCs (paraphrased with 'est une caracteristique' and 'du fait qu'il soit') and two questions target SCs (paraphrased with 'comme par hasard' and 'simplement parce que la plupart'). The final page shows the optional comment box before participants received their completion code.

Instructions:

Bienvenue!

Vous êtes arrivé sur une île étrangère et vous êtes entouré de nombreux objets, créatures et autres objets inconnus. Heureusement, vous avez rencontré des chercheurs qui vous présenteront certaines des choses qui vous entourent, mais ces chercheurs ont également besoin de votre aide pour qualifier plusieurs nouvelles choses. D'abord, ils vous diront ce qu'ils savent sur 30 choses qu'ils ont trouvées sur l'île. Comme les choses qu'ils ont trouvées sont inhabituelles, les chercheurs ont créé un vocabulaire pour les décrire.

Portez une attention toute particulière à ce qu'ils vous disent, car cela vous aidera à réfléchir aux caractéristiques des nouvelles choses que vous et les chercheurs trouverez.

Vous n'avez pas à vous souvenir de l’aspect de chacune des choses. Essayez plutôt de vous faire une idée générale de l’île et des types de caractéristiques que les choses peuvent avoir. Après cela, les chercheurs vous emmèneront dans une autre partie de l'île. Là, on vous demandera ce que vous pensez des caractéristiques des nouvelles choses que vous rencontrez.

→ Click here to continue

progress

Ah! C'est une crogue.

Une crogue a un vitin.

1. Press 1 or click to continue when you are ready.

progress

Si tu regardes là-bas, il y a un clodard.

Les clodards ont une sépite.

1. Press 1 or click to continue when you are ready.

progress

Bon travail! C'était la première étape. Nous allons maintenant passer à une autre région de l'île que les chercheurs n'ont pas encore explorée. Ils espèrent que les caractéristiques qu'ils vous ont données jusqu'à présent seront utiles pour les aider à répondre à certaines questions sur les nouvelles choses qu'ils y trouvent. Les chercheurs vous demanderont de répondre à certaines questions sur ce qu'ils rencontrent en fonction de ce que vous avez appris jusqu'à présent.

Vous ne vous sentirez peut-être pas toujours très sûr de votre réponse, alors essayez de suivre votre première réaction à la question.

Appuyez sur 1 ou sur la barre d'espace pour continuer.

progress

Si tu regardes par là, tu verras une silette et elle a une sépite.

Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une silette?

(Certainement Non)	1	2	3	4	5	6	7	*(Certainement Oui)*

progress

Ceci est un tantour et il a un vitin.

Est-ce que tu penses qu'un tantour, du fait qu'il soit un tantour, a un vitin?

(Certainement Non) 1 2 3 4 5 6 7 *(Certainement Oui)*

progress

Regarde là-bas, il y a un foulain et il a une sépite.

Est-ce que tu penses que les foulains ont, comme par hasard, une sépite?

(Certainement Non) 1 2 3 4 5 6 7 *(Certainement Oui)*

progress

Regarde! Ça c'est un pleutron et il a une sépite.

Est-ce que tu penses que les pleutrons ont une sépite simplement parce que la plupart des pleutrons ont une sépite?

(Certainement Non)	1	2	3	4	5	6	7	*(Certainement Oui)*

OPTIONNEL

Merci pour votre participation. Avant de conclure cette expérience, nous aimerions vous poser une dernière question.

Y a-t-il quelque chose que vous aimeriez dire aux chercheurs ou autre chose que vous avez remarqué pendant l'expérience?

→ Click here to continue

B.2.3 Full list of stimuli

The tables on the following pages include the stimuli for both the training and the testing phase for all the French pseudoword learning studies presented in this thesis.

Stimuli for pseudoword learning studies

Experiment 2: Definite plural vs. indefinite singular subjects

Training stimuli

Voici un blart.	Les blarts ont une sépite.
Voici un blart.	Un blart a une sépite.
Regarde, il y a un molmet.	Les molmets ont une sépite.
Regarde, il y a un molmet.	Un molmet a une sépite.
Si tu regardes là-bas, il y a un clodard.	Les clodards ont une sépite.
Si tu regardes là-bas, il y a un clodard.	Un clodard a une sépite.
Ceci est un melger.	Les melgers ont une sépite.
Ceci est un melger.	Un melger a une sépite.
Par là il y a un blot.	Les blots ont une sépite.
Par là il y a un blot.	Un blot a une sépite.
Ici on a un bronc.	Les broncs ont une sépite.
Ici on a un bronc.	Un bronc a une sépite.
Ah! C'est un grolot.	Les grolots ont une sépite.
Ah! C'est un grolot.	Un grolot a une sépite.
Par ici on a un brabou.	Les brabous ont une sépite.
Par ici on a un brabou.	Un brabou a une sépite.
On a trouvé un pifi.	Les pifis ont une sépite.
On a trouvé un pifi.	Un pifi a une sépite.
Regarde! C'est un nare.	Les nares ont une sépite.
Regarde! C'est un nare.	Un nare a une sépite.
Voici un blenêt.	Les blenêts ont une sépite.
Voici un blenêt.	Un blenêt a une sépite.
Regarde, il y a un dubot.	Les dubots ont une sépite.
Regarde, il y a un dubot.	Un dubot a une sépite.
Si tu regardes là-bas, il y a un tritot.	Les tritots ont une sépite.
Si tu regardes là-bas, il y a un tritot.	Un tritot a une sépite.
Ceci est un babarou.	Les babarous ont une sépite.
Ceci est un babarou.	Un babarou a une sépite.
Par là il y a un jural.	Les jurals ont une sépite.
Par là il y a un jural.	Un jural a une sépite.
Ici on a une puille.	Les puilles ont une sépite.
Ici on a une puille.	Une puille a une sépite.
Ah! C'est une saupe.	Les saupes ont une sépite.
Ah! C'est une saupe.	Une saupe a une sépite.
Par ici on a une glupe.	Les glupes ont une sépite.
Par ici on a une glupe.	Une glupe a une sépite.
On a trouvé une moste.	Les mostes ont une sépite.
On a trouvé une moste.	Une moste a une sépite.
Regarde! C'est une blouque.	Les blouques ont une sépite.

Regarde! C'est une blouque.	Une blouque a une sépite.
Voici une volnette.	Les volnettes ont une sépite.
Voici une volnette.	Une volnette a une sépite.
Regarde, il y a une graze.	Les grazes ont une sépite.
Regarde, il y a une graze.	Une graze a une sépite.
Si tu regardes là-bas, il y a une brebelle.	Les brebelles ont une sépite.
Si tu regardes là-bas, il y a une brebelle.	Une brebelle a une sépite.
Ceci est une crague.	Les cragues ont une sépite.
Ceci est une crague.	Une crague a une sépite.
Par là il y a une malouse.	Les malouses ont une sépite.
Par là il y a une malouse.	Une malouse a une sépite.
Ici on a une frigue.	Les frigues ont une sépite.
Ici on a une frigue.	Une frigue a une sépite.
Ah! C'est une casalle.	Les casalles ont une sépite.
Ah! C'est une casalle.	Une casalle a une sépite.
Par ici on a une croquaille.	Les croquailles ont une sépite.
Par ici on a une croquaille.	Une croquaille a une sépite.
On a trouvé une craise.	Les craises ont une sépite.
On a trouvé une craise.	Une craise a une sépite.
Regarde! C'est une pepesse.	Les pepesses ont une sépite.
Regarde! C'est une pepesse.	Une pepesse a une sépite.
Voici un roufle.	Les roufles ont un vitin.
Voici un roufle.	Un roufle a un vitin.
Regarde, il y a un droplot.	Les droplots ont un vitin.
Regarde, il y a un droplot.	Un droplot a un vitin.
Si tu regardes là-bas, il y a un pomole.	Les pomoles ont un vitin.
Si tu regardes là-bas, il y a un pomole.	Un pomole a un vitin.
Ceci est un lapaud.	Les lapauds ont un vitin.
Ceci est un lapaud.	Un lapaud a un vitin.
Par là il y a un vonore.	Les vonores ont un vitin.
Par là il y a un vonore.	Un vonore a un vitin.
Ici on a un bladon.	Les bladons ont un vitin.
Ici on a un bladon.	Un bladon a un vitin.
Ah! C'est un saleur.	Les saleurs ont un vitin.
Ah! C'est un saleur.	Un saleur a un vitin.
Par ici on a un brion.	Les brions ont un vitin.
Par ici on a un brion.	Un brion a un vitin.
On a trouvé un mionde.	Les miondes ont un vitin.
On a trouvé un mionde.	Un mionde a un vitin.
Regarde! C'est un gonel.	Les gonels ont un vitin.
Regarde! C'est un gonel.	Un gonel a un vitin.
Voici un raster.	Les rasters ont un vitin.
Voici un raster.	Un raster a un vitin.

Regarde, il y a un posparge.	Les posparges ont un vitin.
Regarde, il y a un posparge.	Un posparge a un vitin.
Si tu regardes là-bas, il y a un rouchant.	Les rouchants ont un vitin.
Si tu regardes là-bas, il y a un rouchant.	Un rouchant a un vitin.
Ceci est un cabode.	Les cabodes ont un vitin.
Ceci est un cabode.	Un cabode a un vitin.
Par là il y a un choloir.	Les choloirs ont un vitin.
Par là il y a un choloir.	Un choloir a un vitin.
Ici on a une cripe.	Les cripes ont un vitin.
Ici on a une cripe.	Une cripe a un vitin.
Ah! C'est une mirotte.	Les mirottes ont un vitin.
Ah! C'est une mirotte.	Une mirotte a un vitin.
Par ici on a une prosse.	Les prosses ont un vitin.
Par ici on a une prosse.	Une prosse a un vitin.
On a trouvé une flasse.	Les flasses ont un vitin.
On a trouvé une flasse.	Une flasse a un vitin.
Regarde! C'est une blaque.	Les blaques ont un vitin.
Regarde! C'est une blaque.	Une blaque a un vitin.
Voici une boubelle.	Les boubelles ont un vitin.
Voici une boubelle.	Une boubelle a un vitin.
Regarde, il y a une lattace.	Les lattaces ont un vitin.
Regarde, il y a une lattace.	Une lattace a un vitin.
Si tu regardes là-bas, il y a une rine.	Les rines ont un vitin.
Si tu regardes là-bas, il y a une rine.	Une rine a un vitin.
Ceci est une caie.	Les caies ont un vitin.
Ceci est une caie.	Une caie a un vitin.
Par là il y a une gine.	Les gines ont un vitin.
Par là il y a une gine.	Une gine a un vitin.
Ici on a une sébule.	Les sébules ont un vitin.
Ici on a une sébule.	Une sébule a un vitin.
Ah! C'est une crogue.	Les crogues ont un vitin.
Ah! C'est une crogue.	Une crogue a un vitin.
Par ici on a une tantoute.	Les tantoutes ont un vitin.
Par ici on a une tantoute.	Une tantoute a un vitin.
On a trouvé une prifine.	Les prifines ont un vitin.
On a trouvé une prifine.	Une prifine a un vitin.
Regarde! C'est une distelle.	Les distelles ont un vitin.
Regarde! C'est une distelle.	Une distelle a un vitin.

Training stimuli

Super! On a trouvé un rez et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un rez?
Super! On a trouvé un rez et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un rez?
Super! On a trouvé un rez et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un rez?
Super! On a trouvé un rez et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un rez?
Regarde! Ça c'est un voilot et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un voilot?
Regarde! Ça c'est un voilot et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un voilot?
Regarde! Ça c'est un voilot et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un voilot?
Regarde! Ça c'est un voilot et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un voilot?
Voilà un naché et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un naché?
Voilà un naché et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un naché?
Voilà un naché et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un naché?
Voilà un naché et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un naché?
On cherchait un sauple et voilà, il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un sauple?
On cherchait un sauple et voilà, il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un sauple?
On cherchait un sauple et voilà, il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un sauple?
On cherchait un sauple et voilà, il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un sauple?
Ceci est un ponodin et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un ponodin?
Ceci est un ponodin et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un ponodin?
Ceci est un ponodin et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un ponodin?
Ceci est un ponodin et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un ponodin?
Regarde là-bas, il y a un juvère et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un juvère?
Regarde là-bas, il y a un juvère et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un juvère?
Regarde là-bas, il y a un juvère et il a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être un juvère?
Regarde là-bas, il y a un juvère et il a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être un juvère?

Génial! C'est une nible et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une nible?
Génial! C'est une nible et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une nible?
Génial! C'est une nible et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une nible?
Génial! C'est une nible et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une nible?
On espérait trouver une cholice et voilà, elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une cholice?
On espérait trouver une cholice et voilà, elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une cholice?
On espérait trouver une cholice et voilà, elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une cholice?
On espérait trouver une cholice et voilà, elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une cholice?
Excellent! C'est une golasse et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une golasse?
Excellent! C'est une golasse et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une golasse?
Excellent! C'est une golasse et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une golasse?
Excellent! C'est une golasse et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une golasse?
Si tu regardes par là, tu verras une silette et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une silette?
Si tu regardes par là, tu verras une silette et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une silette?
Si tu regardes par là, tu verras une silette et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une silette?
Si tu regardes par là, tu verras une silette et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une silette?
As-tu vu ça? C'est une latace et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une latace?
As-tu vu ça? C'est une latace et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une latace?
As-tu vu ça? C'est une latace et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une latace?
As-tu vu ça? C'est une latace et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une latace?
Nous avons trouvé une dérave et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une dérave?
Nous avons trouvé une dérave et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une dérave?
Nous avons trouvé une dérave et elle a un vitin.	Est-ce que tu penses qu'avoir un vitin est une caracteristique pour être une dérave?

Nous avons trouvé une dérave et elle a une sépite.	Est-ce que tu penses qu'avoir une sépite est une caracteristique pour être une dérave?
Super! On a trouvé un déri et il a une sépite.	Est-ce que tu penses qu'un déri, du fait qu'il soit un déri, a une sépite?
Super! On a trouvé un déri et il a un vitin.	Est-ce que tu penses qu'un déri, du fait qu'il soit un déri, a un vitin?
Super! On a trouvé un déri et il a un vitin.	Est-ce que tu penses qu'un déri, du fait qu'il soit un déri, a un vitin?
Super! On a trouvé un déri et il a une sépite.	Est-ce que tu penses qu'un déri, du fait qu'il soit un déri, a une sépite?
Regarde! Ça c'est un bogire et il a une sépite.	Est-ce que tu penses qu'un bogire, du fait qu'il soit un bogire, a une sépite?
Regarde! Ça c'est un bogire et il a un vitin.	Est-ce que tu penses qu'un bogire, du fait qu'il soit un bogire, a un vitin?
Regarde! Ça c'est un bogire et il a un vitin.	Est-ce que tu penses qu'un bogire, du fait qu'il soit un bogire, a un vitin?
Regarde! Ça c'est un bogire et il a une sépite.	Est-ce que tu penses qu'un bogire, du fait qu'il soit un bogire, a une sépite?
Voilà un toumileur et il a une sépite.	Est-ce que tu penses qu'un toumileur, du fait qu'il soit un toumileur, a une sépite?
Voilà un toumileur et il a un vitin.	Est-ce que tu penses qu'un toumileur, du fait qu'il soit un toumileur, a un vitin?
Voilà un toumileur et il a un vitin.	Est-ce que tu penses qu'un toumileur, du fait qu'il soit un toumileur, a un vitin?
Voilà un toumileur et il a une sépite.	Est-ce que tu penses qu'un toumileur, du fait qu'il soit un toumileur, a une sépite?
On cherchait un genduit et voilà, il a une sépite.	Est-ce que tu penses qu'un genduit, du fait qu'il soit un genduit, a une sépite?
On cherchait un genduit et voilà, il a un vitin.	Est-ce que tu penses qu'un genduit, du fait qu'il soit un genduit, a un vitin?
On cherchait un genduit et voilà, il a un vitin.	Est-ce que tu penses qu'un genduit, du fait qu'il soit un genduit, a un vitin?
On cherchait un genduit et voilà, il a une sépite.	Est-ce que tu penses qu'un genduit, du fait qu'il soit un genduit, a une sépite?
Ceci est un tantour et il a une sépite.	Est-ce que tu penses qu'un tantour, du fait qu'il soit un tantour, a une sépite?
Ceci est un tantour et il a un vitin.	Est-ce que tu penses qu'un tantour, du fait qu'il soit un tantour, a un vitin?
Ceci est un tantour et il a un vitin.	Est-ce que tu penses qu'un tantour, du fait qu'il soit un tantour, a un vitin?
Ceci est un tantour et il a une sépite.	Est-ce que tu penses qu'un tantour, du fait qu'il soit un tantour, a une sépite?
Regarde là-bas, il y a un daloir et il a une sépite.	Est-ce que tu penses qu'un daloir, du fait qu'il soit un daloir, a une sépite?
Regarde là-bas, il y a un daloir et il a un vitin.	Est-ce que tu penses qu'un daloir, du fait qu'il soit un daloir, a un vitin?
Regarde là-bas, il y a un daloir et il a un vitin.	Est-ce que tu penses qu'un daloir, du fait qu'il soit un daloir, a un vitin?

Regarde là-bas, il y a un daloir et il a une sépite.	Est-ce que tu penses qu'un daloir, du fait qu'il soit un daloir, a une sépite?
Génial! C'est une sébule et elle a une sépite.	Est-ce que tu penses qu'une sébule, du fait qu'elle soit une sébule, a une sépite?
Génial! C'est une sébule et elle a un vitin.	Est-ce que tu penses qu'une sébule, du fait qu'elle soit une sébule, a un vitin?
Génial! C'est une sébule et elle a un vitin.	Est-ce que tu penses qu'une sébule, du fait qu'elle soit une sébule, a un vitin?
Génial! C'est une sébule et elle a une sépite.	Est-ce que tu penses qu'une sébule, du fait qu'elle soit une sébule, a une sépite?
On espérait trouver une fulame et voilà, elle a une sépite.	Est-ce que tu penses qu'une fulame, du fait qu'elle soit une fulame, a une sépite?
On espérait trouver une fulame et voilà, elle a un vitin.	Est-ce que tu penses qu'une fulame, du fait qu'elle soit une fulame, a un vitin?
On espérait trouver une fulame et voilà, elle a un vitin.	Est-ce que tu penses qu'une fulame, du fait qu'elle soit une fulame, a un vitin?
On espérait trouver une fulame et voilà, elle a une sépite.	Est-ce que tu penses qu'une fulame, du fait qu'elle soit une fulame, a une sépite?
Excellent! C'est une faustoque et elle a une sépite.	Est-ce que tu penses qu'une faustoque, du fait qu'elle soit une faustoque, a une sépite?
Excellent! C'est une faustoque et elle a un vitin.	Est-ce que tu penses qu'une faustoque, du fait qu'elle soit une faustoque, a un vitin?
Excellent! C'est une faustoque et elle a un vitin.	Est-ce que tu penses qu'une faustoque, du fait qu'elle soit une faustoque, a un vitin?
Excellent! C'est une faustoque et elle a une sépite.	Est-ce que tu penses qu'une faustoque, du fait qu'elle soit une faustoque, a une sépite?
Si tu regardes par là, tu verras une tifonce et elle a une sépite.	Est-ce que tu penses qu'une tifonce, du fait qu'elle soit une tifonce, a une sépite?
Si tu regardes par là, tu verras une tifonce et elle a un vitin.	Est-ce que tu penses qu'une tifonce, du fait qu'elle soit une tifonce, a un vitin?
Si tu regardes par là, tu verras une tifonce et elle a un vitin.	Est-ce que tu penses qu'une tifonce, du fait qu'elle soit une tifonce, a un vitin?
Si tu regardes par là, tu verras une tifonce et elle a une sépite.	Est-ce que tu penses qu'une tifonce, du fait qu'elle soit une tifonce, a une sépite?
As-tu vu ça? C'est une plavarnie et elle a une sépite.	Est-ce que tu penses qu'une plavarnie, du fait qu'elle soit une plavarnie, a une sépite?
As-tu vu ça? C'est une plavarnie et elle a un vitin.	Est-ce que tu penses qu'une plavarnie, du fait qu'elle soit une plavarnie, a un vitin?
As-tu vu ça? C'est une plavarnie et elle a un vitin.	Est-ce que tu penses qu'une plavarnie, du fait qu'elle soit une plavarnie, a un vitin?
As-tu vu ça? C'est une plavarnie et elle a une sépite.	Est-ce que tu penses qu'une plavarnie, du fait qu'elle soit une plavarnie, a une sépite?
Nous avons trouvé une pescule et elle a une sépite.	Est-ce que tu penses qu'une pescule, du fait qu'elle soit une pescule, a une sépite?
Nous avons trouvé une pescule et elle a un vitin.	Est-ce que tu penses qu'une pescule, du fait qu'elle soit une pescule, a un vitin?
Nous avons trouvé une pescule et elle a un vitin.	Est-ce que tu penses qu'une pescule, du fait qu'elle soit une pescule, a un vitin?

Nous avons trouvé une pescule et elle a une sépite.	Est-ce que tu penses qu'une pescule, du fait qu'elle soit une pescule, a une sépite?
Super! On a trouvé un pauseil et il a une sépite.	Est-ce que tu penses que les pauseils ont une sépite simplement parce que la plupart des pauseils ont une sépite?
Super! On a trouvé un pauseil et il a un vitin.	Est-ce que tu penses que les pauseils ont un vitin simplement parce que la plupart des pauseils ont un vitin?
Super! On a trouvé un pauseil et il a un vitin.	Est-ce que tu penses que les pauseils ont un vitin simplement parce que la plupart des pauseils ont un vitin?
Super! On a trouvé un pauseil et il a une sépite.	Est-ce que tu penses que les pauseils ont une sépite simplement parce que la plupart des pauseils ont une sépite?
Regarde! Ça c'est un pleutron et il a une sépite.	Est-ce que tu penses que les pleutrons ont une sépite simplement parce que la plupart des pleutrons ont une sépite?
Regarde! Ça c'est un pleutron et il a un vitin.	Est-ce que tu penses que les pleutrons ont un vitin simplement parce que la plupart des pleutrons ont un vitin?
Regarde! Ça c'est un pleutron et il a un vitin.	Est-ce que tu penses que les pleutrons ont un vitin simplement parce que la plupart des pleutrons ont un vitin?
Regarde! Ça c'est un pleutron et il a une sépite.	Est-ce que tu penses que les pleutrons ont une sépite simplement parce que la plupart des pleutrons ont une sépite?
Voilà un rôlet et il a une sépite.	Est-ce que tu penses que les rôlets ont une sépite simplement parce que la plupart des rôlets ont une sépite?
Voilà un rôlet et il a un vitin.	Est-ce que tu penses que les rôlets ont un vitin simplement parce que la plupart des rôlets ont un vitin?
Voilà un rôlet et il a un vitin.	Est-ce que tu penses que les rôlets ont un vitin simplement parce que la plupart des rôlets ont un vitin?
Voilà un rôlet et il a une sépite.	Est-ce que tu penses que les rôlets ont une sépite simplement parce que la plupart des rôlets ont une sépite?
On cherchait un tupent et voilà, il a une sépite.	Est-ce que tu penses que les tupents ont une sépite simplement parce que la plupart des tupents ont une sépite?
On cherchait un tupent et voilà, il a un vitin.	Est-ce que tu penses que les tupents ont un vitin simplement parce que la plupart des tupents ont un vitin?
On cherchait un tupent et voilà, il a un vitin.	Est-ce que tu penses que les tupents ont un vitin simplement parce que la plupart des tupents ont un vitin?
On cherchait un tupent et voilà, il a une sépite.	Est-ce que tu penses que les tupents ont une sépite simplement parce que la plupart des tupents ont une sépite?
Ceci est un troumet et il a une sépite.	Est-ce que tu penses que les troumets ont une sépite simplement parce que la plupart des troumets ont une sépite?
Ceci est un troumet et il a un vitin.	Est-ce que tu penses que les troumets ont un vitin simplement parce que la plupart des troumets ont un vitin?

Ceci est un troumet et il a un vitin.	Est-ce que tu penses que les troumets ont un vitin simplement parce que la plupart des troumets ont un vitin?
Ceci est un troumet et il a une sépite.	Est-ce que tu penses que les troumets ont une sépite simplement parce que la plupart des troumets ont une sépite?
Regarde là-bas, il y a un borcier et il a une sépite.	Est-ce que tu penses que les borciers ont une sépite simplement parce que la plupart des borciers ont une sépite?
Regarde là-bas, il y a un borcier et il a un vitin.	Est-ce que tu penses que les borciers ont un vitin simplement parce que la plupart des borciers ont un vitin?
Regarde là-bas, il y a un borcier et il a un vitin.	Est-ce que tu penses que les borciers ont un vitin simplement parce que la plupart des borciers ont un vitin?
Regarde là-bas, il y a un borcier et il a une sépite.	Est-ce que tu penses que les borciers ont une sépite simplement parce que la plupart des borciers ont une sépite?
Génial! C'est une porite et elle a une sépite.	Est-ce que tu penses que les porites ont une sépite simplement parce que la plupart des porites ont une sépite?
Génial! C'est une porite et elle a un vitin.	Est-ce que tu penses que les porites ont un vitin simplement parce que la plupart des porites ont un vitin?
Génial! C'est une porite et elle a un vitin.	Est-ce que tu penses que les porites ont un vitin simplement parce que la plupart des porites ont un vitin?
Génial! C'est une porite et elle a une sépite.	Est-ce que tu penses que les porites ont une sépite simplement parce que la plupart des porites ont une sépite?
On espérait trouver une siboutie et voilà, elle a une sépite.	Est-ce que tu penses que les sibouties ont une sépite simplement parce que la plupart des sibouties ont une sépite?
On espérait trouver une siboutie et voilà, elle a un vitin.	Est-ce que tu penses que les sibouties ont un vitin simplement parce que la plupart des sibouties ont un vitin?
On espérait trouver une siboutie et voilà, elle a un vitin.	Est-ce que tu penses que les sibouties ont un vitin simplement parce que la plupart des sibouties ont un vitin?
On espérait trouver une siboutie et voilà, elle a une sépite.	Est-ce que tu penses que les sibouties ont une sépite simplement parce que la plupart des sibouties ont une sépite?
Excellent! C'est une gascute et elle a une sépite.	Est-ce que tu penses que les gascutes ont une sépite simplement parce que la plupart des gascutes ont une sépite?
Excellent! C'est une gascute et elle a un vitin.	Est-ce que tu penses que les gascutes ont un vitin simplement parce que la plupart des gascutes ont un vitin?
Excellent! C'est une gascute et elle a un vitin.	Est-ce que tu penses que les gascutes ont un vitin simplement parce que la plupart des gascutes ont un vitin?

Excellent! C'est une gascute et elle a une sépite.	Est-ce que tu penses que les gascutes ont une sépite simplement parce que la plupart des gascutes ont une sépite?
Si tu regardes par là, tu verras une celtire et elle a une sépite.	Est-ce que tu penses que les celtires ont une sépite simplement parce que la plupart des celtires ont une sépite?
Si tu regardes par là, tu verras une celtire et elle a un vitin.	Est-ce que tu penses que les celtires ont un vitin simplement parce que la plupart des celtires ont un vitin?
Si tu regardes par là, tu verras une celtire et elle a un vitin.	Est-ce que tu penses que les celtires ont un vitin simplement parce que la plupart des celtires ont un vitin?
Si tu regardes par là, tu verras une celtire et elle a une sépite.	Est-ce que tu penses que les celtires ont une sépite simplement parce que la plupart des celtires ont une sépite?
As-tu vu ça? C'est une clune et elle a une sépite.	Est-ce que tu penses que les clunes ont une sépite simplement parce que la plupart des clunes ont une sépite?
As-tu vu ça? C'est une clune et elle a un vitin.	Est-ce que tu penses que les clunes ont un vitin simplement parce que la plupart des clunes ont un vitin?
As-tu vu ça? C'est une clune et elle a un vitin.	Est-ce que tu penses que les clunes ont un vitin simplement parce que la plupart des clunes ont un vitin?
As-tu vu ça? C'est une clune et elle a une sépite.	Est-ce que tu penses que les clunes ont une sépite simplement parce que la plupart des clunes ont une sépite?
Nous avons trouvé une sembe et elle a une sépite.	Est-ce que tu penses que les sembes ont une sépite simplement parce que la plupart des sembes ont une sépite?
Nous avons trouvé une sembe et elle a un vitin.	Est-ce que tu penses que les sembes ont un vitin simplement parce que la plupart des sembes ont un vitin?
Nous avons trouvé une sembe et elle a un vitin.	Est-ce que tu penses que les sembes ont un vitin simplement parce que la plupart des sembes ont un vitin?
Nous avons trouvé une sembe et elle a une sépite.	Est-ce que tu penses que les sembes ont une sépite simplement parce que la plupart des sembes ont une sépite?
Super! On a trouvé un sersalet et il a une sépite.	Est-ce que tu penses que les sersalets ont, comme par hasard, une sépite?
Super! On a trouvé un sersalet et il a un vitin.	Est-ce que tu penses que les sersalets ont, comme par hasard, un vitin?
Super! On a trouvé un sersalet et il a un vitin.	Est-ce que tu penses que les sersalets ont, comme par hasard, un vitin?
Super! On a trouvé un sersalet et il a une sépite.	Est-ce que tu penses que les sersalets ont, comme par hasard, une sépite?
Regarde! Ça c'est un suglé et il a une sépite.	Est-ce que tu penses que les suglés ont, comme par hasard, une sépite?
Regarde! Ça c'est un suglé et il a un vitin.	Est-ce que tu penses que les suglés ont, comme par hasard, un vitin?
Regarde! Ça c'est un suglé et il a un vitin.	Est-ce que tu penses que les suglés ont, comme par hasard, un vitin?
Regarde! Ça c'est un suglé et il a une sépite.	Est-ce que tu penses que les suglés ont, comme par hasard, une sépite?

Voilà un bobomour et il a une sépite.	Est-ce que tu penses que les bobomours ont, comme par hasard, une sépite?
Voilà un bobomour et il a un vitin.	Est-ce que tu penses que les bobomours ont, comme par hasard, un vitin?
Voilà un bobomour et il a un vitin.	Est-ce que tu penses que les bobomours ont, comme par hasard, un vitin?
Voilà un bobomour et il a une sépite.	Est-ce que tu penses que les bobomours ont, comme par hasard, une sépite?
On cherchait un croilet et voilà, il a une sépite.	Est-ce que tu penses que les croilets ont, comme par hasard, une sépite?
On cherchait un croilet et voilà, il a un vitin.	Est-ce que tu penses que les croilets ont, comme par hasard, un vitin?
On cherchait un croilet et voilà, il a un vitin.	Est-ce que tu penses que les croilets ont, comme par hasard, un vitin?
On cherchait un croilet et voilà, il a une sépite.	Est-ce que tu penses que les croilets ont, comme par hasard, une sépite?
Ceci est un diplant et il a une sépite.	Est-ce que tu penses que les diplants ont, comme par hasard, une sépite?
Ceci est un diplant et il a un vitin.	Est-ce que tu penses que les diplants ont, comme par hasard, un vitin?
Ceci est un diplant et il a un vitin.	Est-ce que tu penses que les diplants ont, comme par hasard, un vitin?
Ceci est un diplant et il a une sépite.	Est-ce que tu penses que les diplants ont, comme par hasard, une sépite?
Regarde là-bas, il y a un foulain et il a une sépite.	Est-ce que tu penses que les foulains ont, comme par hasard, une sépite?
Regarde là-bas, il y a un foulain et il a un vitin.	Est-ce que tu penses que les foulains ont, comme par hasard, un vitin?
Regarde là-bas, il y a un foulain et il a un vitin.	Est-ce que tu penses que les foulains ont, comme par hasard, un vitin?
Regarde là-bas, il y a un foulain et il a une sépite.	Est-ce que tu penses que les foulains ont, comme par hasard, une sépite?
Génial! C'est une toucote et elle a une sépite.	Est-ce que tu penses que les toucotes ont, comme par hasard, une sépite?
Génial! C'est une toucote et elle a un vitin.	Est-ce que tu penses que les toucotes ont, comme par hasard, un vitin?
Génial! C'est une toucote et elle a un vitin.	Est-ce que tu penses que les toucotes ont, comme par hasard, un vitin?
Génial! C'est une toucote et elle a une sépite.	Est-ce que tu penses que les toucotes ont, comme par hasard, une sépite?
On espérait trouver une prérime et voilà, elle a une sépite.	Est-ce que tu penses que les prérimes ont, comme par hasard, une sépite?
On espérait trouver une prérime et voilà, elle a un vitin.	Est-ce que tu penses que les prérimes ont, comme par hasard, un vitin?
On espérait trouver une prérime et voilà, elle a un vitin.	Est-ce que tu penses que les prérimes ont, comme par hasard, un vitin?

On espérait trouver une prérime et voilà, elle a une sépite.	Est-ce que tu penses que les prérimes ont, comme par hasard, une sépite?
Excellent! C'est une rapsique et elle a une sépite.	Est-ce que tu penses que les rapsiques ont, comme par hasard, une sépite?
Excellent! C'est une rapsique et elle a un vitin.	Est-ce que tu penses que les rapsiques ont, comme par hasard, un vitin?
Excellent! C'est une rapsique et elle a un vitin.	Est-ce que tu penses que les rapsiques ont, comme par hasard, un vitin?
Excellent! C'est une rapsique et elle a une sépite.	Est-ce que tu penses que les rapsiques ont, comme par hasard, une sépite?
Si tu regardes par là, tu verras une mopale et elle a une sépite.	Est-ce que tu penses que les mopales ont, comme par hasard, une sépite?
Si tu regardes par là, tu verras une mopale et elle a un vitin.	Est-ce que tu penses que les mopales ont, comme par hasard, un vitin?
Si tu regardes par là, tu verras une mopale et elle a un vitin.	Est-ce que tu penses que les mopales ont, comme par hasard, un vitin?
Si tu regardes par là, tu verras une mopale et elle a une sépite.	Est-ce que tu penses que les mopales ont, comme par hasard, une sépite?
As-tu vu ça? C'est une crisonne et elle a une sépite.	Est-ce que tu penses que les crisonnes ont, comme par hasard, une sépite?
As-tu vu ça? C'est une crisonne et elle a un vitin.	Est-ce que tu penses que les crisonnes ont, comme par hasard, un vitin?
As-tu vu ça? C'est une crisonne et elle a un vitin.	Est-ce que tu penses que les crisonnes ont, comme par hasard, un vitin?
As-tu vu ça? C'est une crisonne et elle a une sépite.	Est-ce que tu penses que les crisonnes ont, comme par hasard, une sépite?
Nous avons trouvé une phonine et elle a une sépite.	Est-ce que tu penses que les phonines ont, comme par hasard, une sépite?
Nous avons trouvé une phonine et elle a un vitin.	Est-ce que tu penses que les phonines ont, comme par hasard, un vitin?
Nous avons trouvé une phonine et elle a un vitin.	Est-ce que tu penses que les phonines ont, comme par hasard, un vitin?
Nous avons trouvé une phonine et elle a une sépite.	Est-ce que tu penses que les phonines ont, comme par hasard, une sépite?

C
Stimuli for Kevtas experiments

Contents

C.1 Participant Information Sheet and consent form

C.1.1 Consent (adult participants)

All MTurk workers provided informed consent and were given the following Participant Information Sheet. The same, or an updated version (e.g. for GDPR regulations) was provided for all subsequent experiment, whether hosted on Amazon's MTurk or Prolific/Qualtrics. These were all approved under the same ethics application (CUREC R55398-RE001 through R55398-RE004).

Do you know about kevtas?

General Information

The aim of this study is to better understand how language influences how we think about the world.

We appreciate your interest in participating in this online survey. You have been invited to participate, as you are an adult, native speaker of English with no history of a language disorder. Please read through these terms before agreeing to participate by ticking the 'yes' box below. You may ask any questions before taking part by contacting the researcher (details below).

We, the University of Oxford in collaboration with the University of Michigan, are investigating the role of language in the acquisition of concepts to find out more about language that refers to "kinds of things".

You will see some sentences and pictures, and then answer questions on what you have seen. It should take about 5 minutes. No background knowledge is required. The information will be used by trained researchers at the University of Oxford and the University of Michigan. No personally identifiable information will be used or stored with your answers.

Do I have to take part?

Please note that your participation is voluntary. You may withdraw at any point during the questionnaire for any reason, before submitting your answers, by pressing the closing the browser. However, we are only able to reimburse participants who complete the full survey.

How will your data be used?

Your answers will be completely anonymous, and we will use all reasonable endeavours to keep them confidential.

Your data will be stored in a password-protected file and may be used in academic publications. Your IP address will not be stored and 'Requesters' on Amazon's Mechanical Turk are not allowed to access any personally identifiable information. 'Requesters' only have access to the MTurk Worker ID for purposes of remuneration. All questions are optional. Research data will be stored for a minimum of three years after publication or public release.

The data that we collect from you may be transferred to, and stored or processed at, a destination outside your country. By submitting your personal data, you agree to this transfer, storing or processing.

Who will have access to your data?

Amazon's Mechanical Turk is the data controller with respect to your personal data and, as such, will determine how your personal data is used. Please see their privacy notice here [https://www.amazon.com/gp/help/customer/display.html/ref=footer_privacy?ie=UTF8&nodeId=468496]. Amazon's Mechanical Turk will share only fully anonymised data with the University of Oxford and the University of Michigan, for the purposes of research.

Your information may be shared with other fully trained researchers at the University of Oxford and the University of Michigan.

Responsible members of the University of Oxford and the University of Michigan, funders may be given access to data for monitoring and/or audit of the study to ensure we are complying with guidelines, or as otherwise required by law.

This questionnaire is for a DPhil/PhD project. The principal researcher is Kim Fuellenbach, who is attached to the Linguistics Department at the University of Oxford. This project is being completed under the supervision of Dr. E. Matthew Husband (University of Oxford) and Dr. Susan Gelman (University of Michigan).

This project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee [CUREC reference number R55398_RE004] and the University of Michigan IRB.

What if there is a problem?

If you have a concern about any aspect of this project, please speak to the researcher [Kim Fuellenbach, kim.fuellenbach@ling-phil.ox.ac.uk] or their supervisor [Susan Gelman, gelman@umich.edu or Matt Husband, matthew.husband@ling-phil.ox.ac.uk], who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the relevant Chair of the Research Ethics Committee at the University of Oxford:

Chair, Social Sciences & Humanities Interdivisional Research Ethics Committee; Email: ethics@socsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

The Chair will seek to resolve the matter in a reasonably expeditious manner.

Please note that you may only participate in this survey if you are 18 years of age or over.

□ I certify that I am 18 years of age or over.

If you have read the information above and agree to participate with the understanding that the data (including any personal data) you submit will be processed accordingly, please check the relevant box below to get started.

□ Yes, I agree to take part

C.1.2 Consent (child participants)

For all studies including child participants, run in the Conceptual Development Lab at the University of Michigan or in the Living Labs in Ann Arbor, all parents provided consent for their child's participation.

Consent to Participate in Psychology Study

Study Name: Do you know about kevtas?

This study is being run by Susan Gelman, a researcher at the University of Michigan and Kim Fuellenbach, a researcher at the University of Oxford.

Invitation to Participate. With your permission, your child will participate in this research project. We are very interested in how children learn about the world. Therefore, we invite you to fill out this questionnaire.

Risks of Participation. There are no known risks associated with participation in this questionnaire. If you are uncomfortable for any reason and wish to end participation, you are free to stop at any time. To stop, simply return the questionnaire and tell the researcher you decided not to fill it out. We will then shred the questionnaire.

Benefits of Participation. There is no direct benefit to you that you can expect to receive as a result of participating in this study. However, we ultimately hope that our findings can add in helpful ways to what we know about ideas about gender.

Costs

Your participation in this study involves no cost to you.

Confidentiality

We hope to publish the results of this study, but will not include information that identifies you. There are reasons why people other than the researchers may ask to see information you provided as part of the study. This includes the University of Michigan, which is responsible for making sure the research is done safely and properly. To keep your information safe, the researchers will store consent forms separately from data in a secure location. The data provided during the experiment will be assigned an ID number that will be not be stored with your identity or name. Links between names and ID numbers will be destroyed after data collection is complete within one year from the date of your participation, to further protect confidentiality.

Storage and future use of data

The data will be stored in a locked and secure location at the University of Michigan indefinitely. We might share that data with other researchers outside the University. However, we will not share your name or other information that could identify you.

Voluntary nature of the study

Participating in this study is voluntary. Even if you decide to participate now, you may change your mind and stop at any time. To stop, simply return the questionnaire and tell the researcher you decided not to fill it out. If you decide to withdraw, your data will not be retained. You will be able to withdraw from the study up to one week after participation.

Contact Information.

If you have questions about this research, you may contact the principal investigator: Dr. Susan Gelman (gelman@umich.edu; 734-764- 0268); Mail can be sent to the investigator at: 530 Church St., Department of Psychology, Ann Arbor MI 48109. Alternatively, you may contact the researcher: Kim Fuellenbach (kim.fuellenbach@ling-phil.ox.ac.uk).

If you have questions about your rights as a research participant, or wish to discuss any concerns about this study with someone other than the researcher(s), please contact the University of Michigan Institutional Review Board, 2800 Plymouth Road, Building 520, Room 1169, Ann Arbor, MI 48109-2800, (734) 936-0933 [or toll free, (866) 936-0933], irbhsbs@umich.edu.

If you understand the above points and choose to participate in this study, please sign below:

____ ____________________

Your Name **Date**

C.1.3 Child assent

All children provided oral assent for their own participation.

GTS Study Child Assent

Chat with child, e.g. ask them how their day is going, interact with them while they play with the toys, etc.

When parent has completed consent:

Great, I see that you have agreed to participate. Do you have any questions for me? And just as a reminder, we will be video taping the whole session.

To child:
[Child name], Your parent/mom/dad said it is OK to play the game with me. Would you like to play a game and answer some questions with me? You would? Great! Let me show you the other room with the toys.

[Escort parent to the viewing room (or into the study room if the child prefers the parent to be with them); Escort child into the testing room]

We're ready to get started with our game. Are you ready? … Great!

C.1.4 Living Lab Amendment

For carrying out the *kevtas* study through the Living Lab initiative in the Museum of Natural History in Ann Arbor, the consent form for child participation was adapted as follows.

Consent for Child's Participation in a Research Study
Study Name: Do you know about kevtas?

This study is being run by Susan Gelman, a researcher at the University of Michigan, and Kim Fuellenbach, a researcher at the University of Oxford.

Invitation to participate in a study
We invite your child to participate in a study that explores how children interpret “generic language" (that is, statements that refer to general categories).

Description of the study procedure
The study lasts approximately ten minutes. If you agree to let your child participate, a trained researcher will read a book with your child. The picture book is about novel, made-up kinds of animals. On each page of the book, there will be a picture of one of these animals and a description of a distinctive property that pertains to the animal. We then present two more pictures of novel animals, similar to the first one and ask your child a question about these. Additionally, we may ask your child to participate in a series of tasks that will examine comprehension, memory, vocabulary, language, and beliefs. These tasks are designed to help us better understand how children learn about the world, think about new animals and reason about categories.

Benefits
There are no direct benefits to participants, but children enjoy looking at pictures and sharing their knowledge with a researcher. Additionally, others may ultimately benefit from the knowledge obtained in this research. We do not evaluate individual children’s performance. Rather, we are interested in children as a group in order to discover how individuals of different ages develop.

Costs and Payments
Your participation in this study involves no cost to you. Your child will receive a small token gift (e.g., a small toy), even if you, or your child, choose to withdraw from the study.

Risks and discomforts
There are no known risks associated with this study; however, breach of confidentiality is one possible risk. When collecting data, identifying information will only be provided on consent forms, and consent forms will be stored separately from data. Only authorized study team members will have access to forms and data. In addition, if your child feels bored or worries that there are right and wrong answers, the researcher will stop the study. Also, if at any time your child indicates he/she is unwilling to participate, we will stop the session.

Confidentiality
We hope to publish the results of this study, but will not include information that identifies you or your child. There are reasons why people other than the researchers may ask to see information you and your child provided as part of the study. This includes the University of Michigan, which is responsible for making sure the research is done safely and properly. To keep your information safe, the researchers will store consent forms separately from the collected data. The data provided during the experiment will be assigned an ID number that will be not be stored with your identity or name. Links between names and ID numbers will be destroyed after data collection is complete, to further protect confidentiality.

Storage and future use of data
The paper data will be stored in a locked and secure location at the University of Michigan indefinitely. We might share that data with other researchers outside the University, only if you provide your consent to that below. However, we will not share your name or other information that could identify you or your child.

Voluntary nature of the study
Participating in this study is voluntary. Even if you decide to participate now, you may change your mind and

stop at any time. If your child decides to withdraw, their data will not be retained or saved digitally.

Contact information

If you have questions about this research, you may contact the principal investigator, Susan Gelman (gelman@umich.edu). Mail can be sent to Susan Gelman at: 530 Church St., Department of Psychology, Ann Arbor MI 48109. If you have questions about your rights as a research participant, or wish to discuss any concerns about this study with someone other than the researchers, please contact the University of Michigan Institutional Review Board, 2800 Plymouth Rd., Building 520, Room 1169, Ann Arbor, MI 48109- 2800, (734) 936-0933 [or toll free, (866) 936-0933], irbhsbs@umich.edu.

Parental Permission

By signing this document, you are indicating that you are aware of the points made above, that you are agreeing to allow your child to be part of this study. If you would like, you will be given a copy of this document for your records, and one copy will be kept with the study records. Be sure that the questions you have asked about the study have been answered and that you understand what your child will be asked to do. You may contact the researcher if you think of a question later.

________________________	________________________	__________
Print Parent Name	*Signature of Parent*	*Date*
________________________	________________________	M F
Print Child Name	*Child DOB (MM/DD/YY)*	*Child Gender (circle)*

____

Address *City* *State* *Zip*

C.2 Demographics questionnaire

Adult participants and parents of child participants were asked to fill out a combined sociolinguistics and demographics questionnaire. However, all questions in this questionnaire were optional.

Optional Demographics

Optional Demographics Section

The researchers would like to know a little bit about you. Please fill in the following part.
Note: This is an optional section.

Basic Information

What is your age? ☐

Race/Ethnicity (Check all that apply)

- ☐ White
- ☐ Latino or Hispanic
- ☐ Black or African American
- ☐ Asian or Asian American
- ☐ American Indian or Alaska Native
- ☐ Middle Eastern or North African
- ☐ Native Hawaiian or Other Pacific Islander
- ☐ Not Listed

☐

With which gender do you most identify?

- Man
- Woman
- Transgender Man
- Transgender Woman
- Gender Non-Conforming
- Not Listed
- Prefer Not to Answer

What is your native language?

- English
- non-English

Did you live in an English-speaking environment from birth till (at least) age 13?

- Yes
- No

Did both of your parents speak English to you at home?

- Yes
- No

→

C.3 Essentialism questionnaire

Adult participants and parents of child participants were asked to complete an essentialism questionnaire. The questions were a subset of the longer essentialism questionnaire employed in the *zarpies* studies.

Optional Demographics Questionnaire

This part will ask you a little bit about your beliefs. Please respond to the following ten statements, rating them on a scale from Strongly Disagree to Strongly Agree.

People who are mathematically talented at age 20 tend to be mathematically talented at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

People who are athletic at age 20 tend to be athletic at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

People who are intelligent at age 20 tend to be intelligent at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

People who are goal-oriented at age 20 tend to be goal-oriented at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

People who are shy at age 20 tend to be shy at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

People who are stubborn at age 20 tend to be stubborn at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

People who are humorous at age 20 tend to be humorous at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

People who are cautious at age 20 tend to be cautious at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

People who are artistic at age 20 tend to be artistic at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

People who are pessimistic at age 20 tend to be pessimistic at age 40.

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Please rate	○	○	○	○	○	○

→

C.4 Stimuli baseline test

This section includes a complete trial of the baseline studies to ascertain the subset of stimuli used for the *kevtas* studies. They are similar to the full version, presented in Section C.5 below but do not include any information about principally or statistically connected properties, neither in generic not in specific language.

Note that the full trial, including the two options (shape and property match) only became available after the audio was played. This incremental presentation is visually presented for the first trial here but, for reasons of space, not thereafter.

0% 100%

Welcome!
In this task you will be seeing pictures of some new animals and hearing new facts. These are animals are not real animals. These are new animals you never saw before, so you will be learning some new names.

→

0% 100%

Today we're going to do a short activity and I am going to ask you some short questions. There are no right or wrong answers so you can say and do whatever you think is right at the time. You may stop at any time during our talk, even if we're not done yet.

Do you want to give it a try?

→

0% 100%

ID Number

Where is this located?

- ◯ Conceptual Development Lab
- ◯ Natural History Museum
- ◯ Handson Museum

CHILD AGE

- ◯ 4-5
- ◯ 6-7
- ◯ 8-10

CHI GENDER

- ◯ M
- ◯ F

0% 100%

Welcome!

In this task you will be seeing pictures of some new animals and hearing new facts. These are animals are not real animals. These are new animals you never saw before, so you will be learning some new names.

→

0% 100%

Today we're going to do a short activity and I am going to ask you some short questions. There are no right or wrong answers so you can say and do whatever you think is right at the time. You may stop at any time during our talk, even if we're not done yet.

Do you want to give it a try?

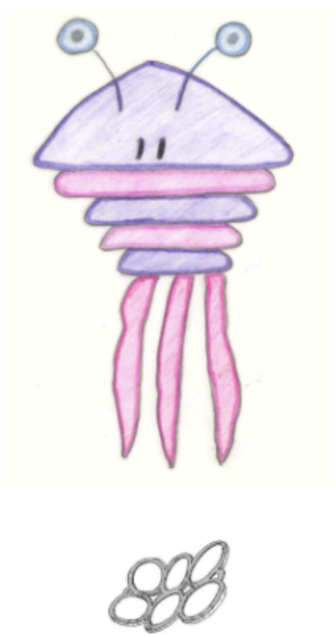
→

0% 100%

Which one of these is also a bactra?

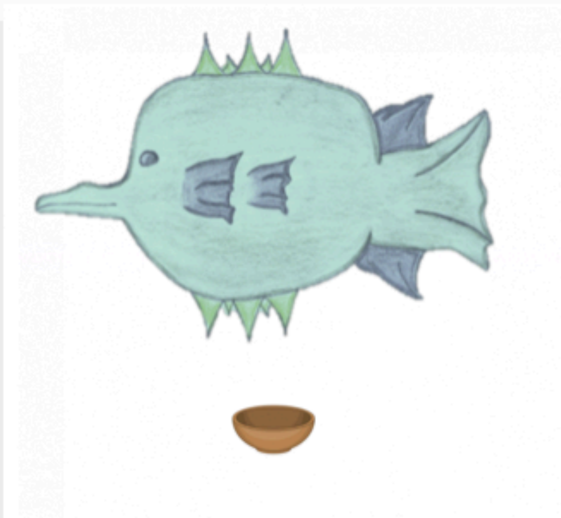
→

0% 100%

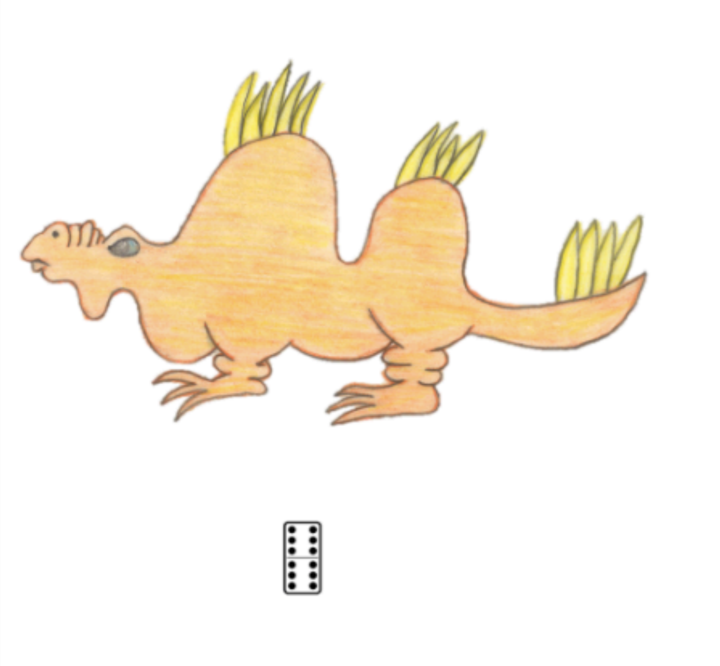


Which one of these is also a bactra?

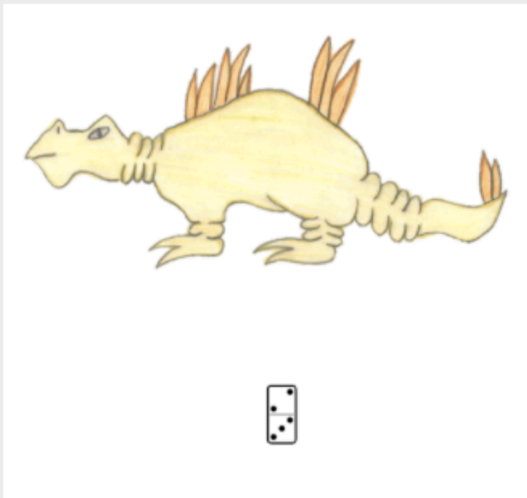
Which one of these is also a bant?



0% 100%

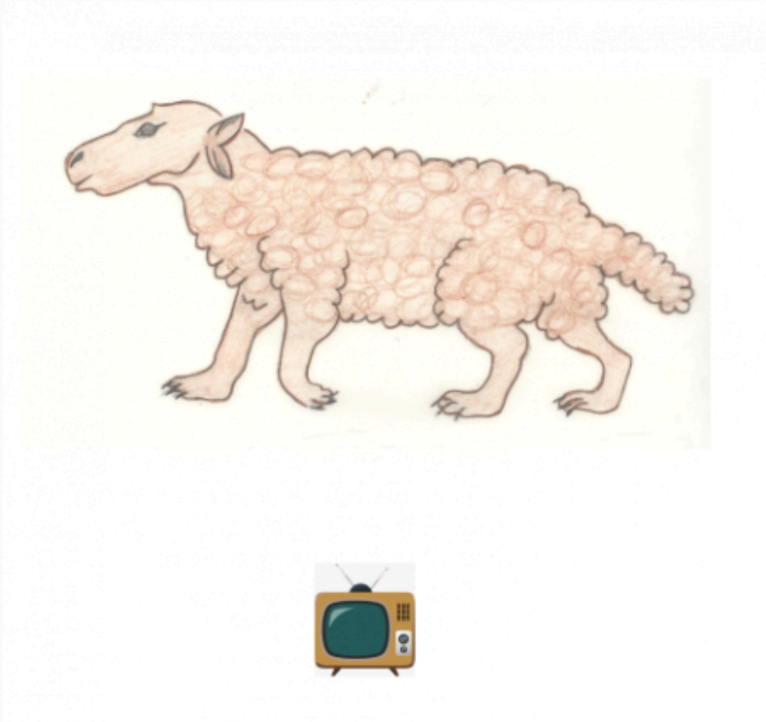


Which one of these is also a febbit?

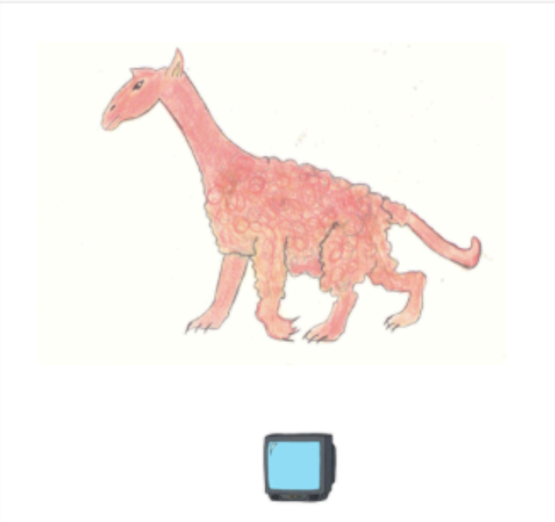


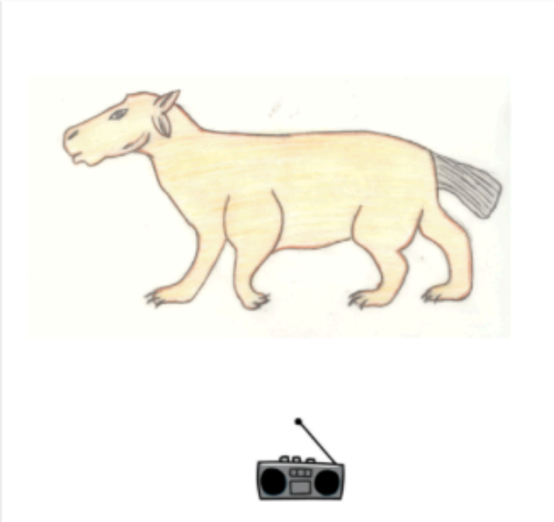
0% 100%

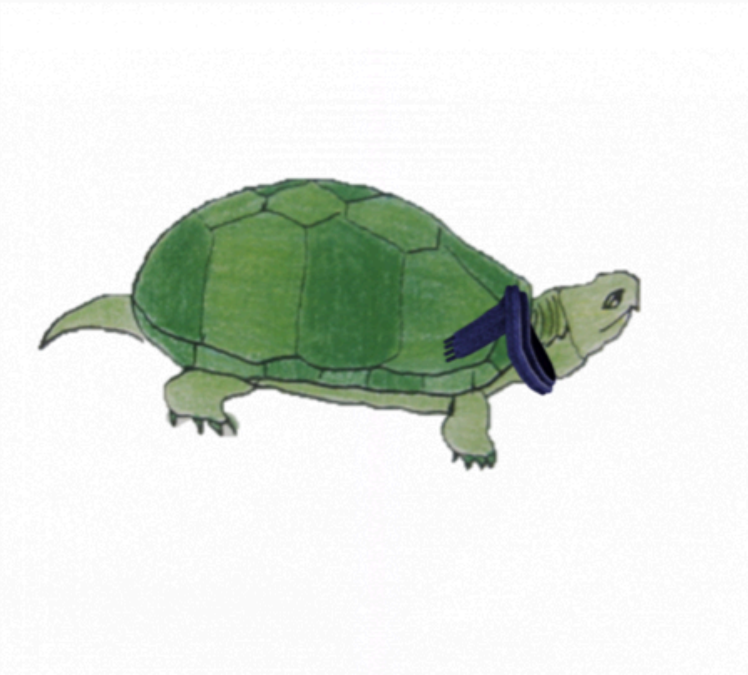
Which one of these is also a fep?



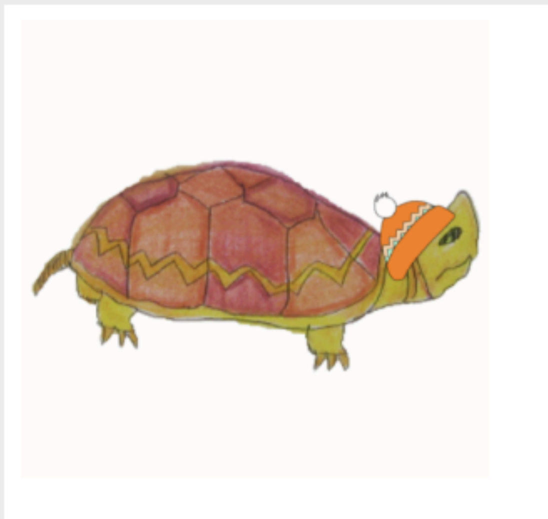
Which one of these is also a kevta?







Which one of these is also a plog?



Which one of these is also a scred?

Which one of these is also a vorzyd?

C.5 Stimuli

Below is a complete trial of the *kevtas* studies that were run online (in the BP wording condition). This study was hosted on Qualtrics and was designed for adult participants. Note that this is only one of the possible randomised trials. Further note that this study did not include the full essentialism questionnaire but the reduced version.

For the studies that included audio stimuli, there was a small audio button at the top of the target image. The audio stimuli are available separately.

Consent to Participate in Psychology Study
This study is being conducted by Susan Gelman and Kim Fuellenbach, researchers at the University of Michigan and University of Oxford, respectively.

Invitation to Participate in Our Study
You are invited because you are a native speaker of English. You are being asked to take part in a research study which explores how children and adults think about the world. We invite you to fill out this questionnaire.

Study Procedure
The study will take around 10 minutes to complete, and involves taking a survey on a computer. You will answer some questions about animals.

Risks of Participation
The risks of participation in this online study are those associated with basic computer tasks, boredom, fatigue, mild stress or breach of confidentiality. If you are uncomfortable for any reason and wish to end participation, you are free to stop at any time. To stop, simply close the browser window.

Benefits of Participation
There is no direct benefit to you that you can expect to receive as a result of participating in this study. However, we ultimately hope that our findings can add in helpful ways to what we know about child development.

Confidentiality of Records
We hope to publish the results of this study but will not include information that identifies you. Your Mechanical Turk Worker ID will be stored separately from the data we collect

Confidentiality of Records

We hope to publish the results of this study but will not include information that identifies you. Your Mechanical Turk Worker ID will be stored separately from the data we collect from you. Please be aware that your MTurk Worker ID can be potentially linked to information about you on your Amazon public profile page, depending on settings you have for you Amazon profile. We will not be accessing any personally identifying information about you that you may have put on your Amazon public profile page.

Voluntary Participation

Your participation in this study is voluntary. You are to stop participating at any time by closing the browser window. In the event that you close your browser during the study, the information you have already provided will not be analyzed.

Do you consent to participate in our study?

Yes, I acknowledge that I'm at least 18, I am a native speaker of English and consent to participate in this research study.

Click to write the question text

I'm not a robot

reCAPTCHA
Privacy - Terms

→

Welcome!

In this task, you will be seeing pictures of some new animals. They are different from the animals we know, so you will learn the name for each of them. For each animal, you will see a description and a small picture of something in the context.

Then you will be asked to identify another one. For this, we will show you two other pictures to choose from.

Please answer to the best of your ability.

When you're ready, please turn off any distractions, such as your cell phone, for the duration of the experiment! Please avoid using the mobile version of this site.

Timing

These page timer metrics will not be displayed to the recipient.

First Click	0 seconds
Last Click	0 seconds
Page Submit	0 seconds
Click Count	0 clicks

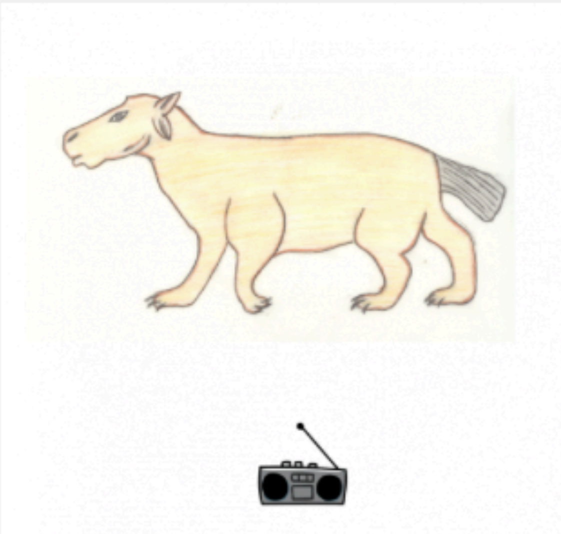
Feps wear collars.

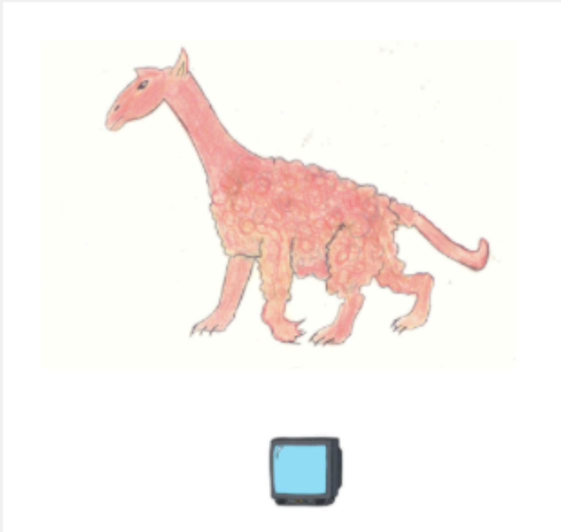
Which of these is also a fep?

→

Kevtas watch TV.

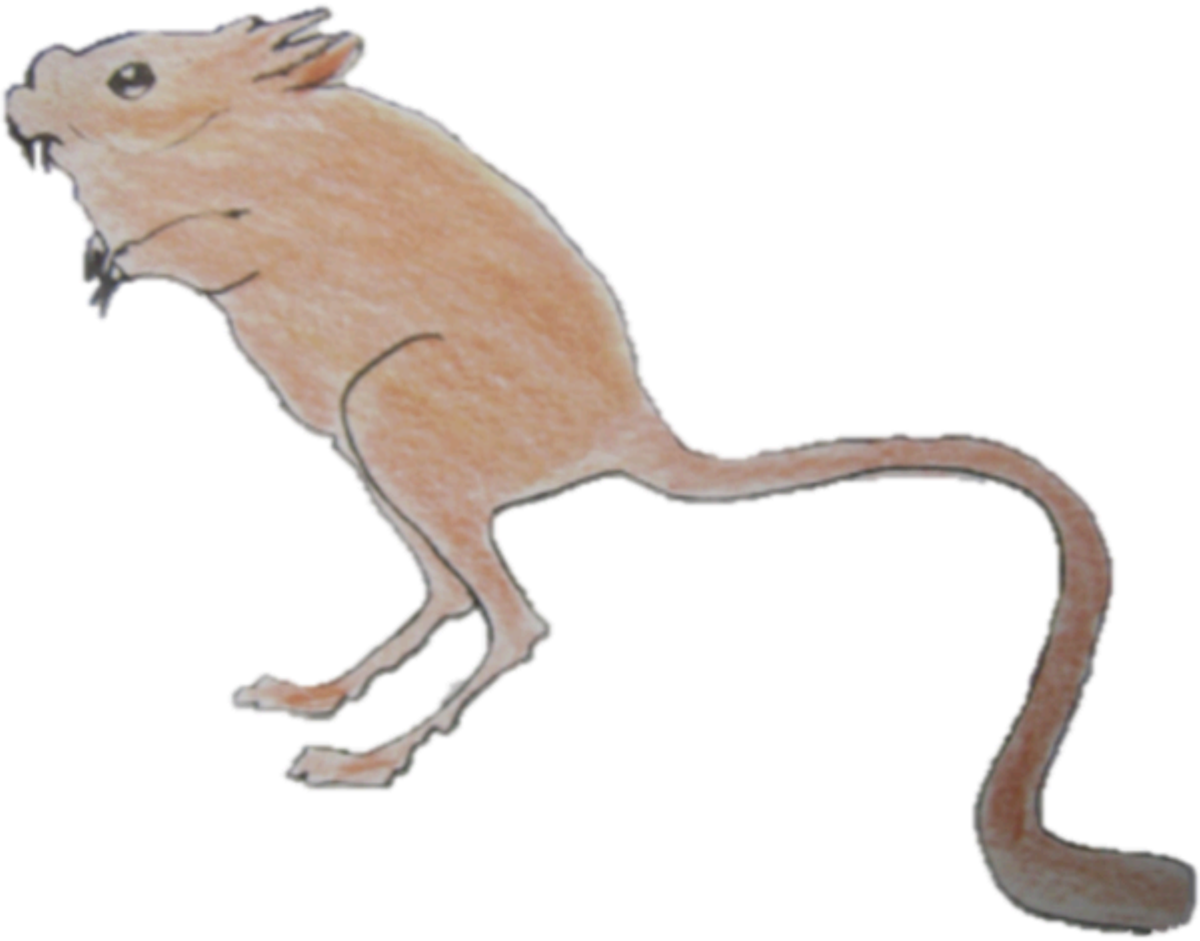
Which of these is also a kevta?





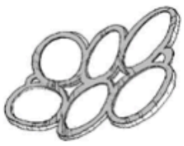
→

Screds sit on couches.

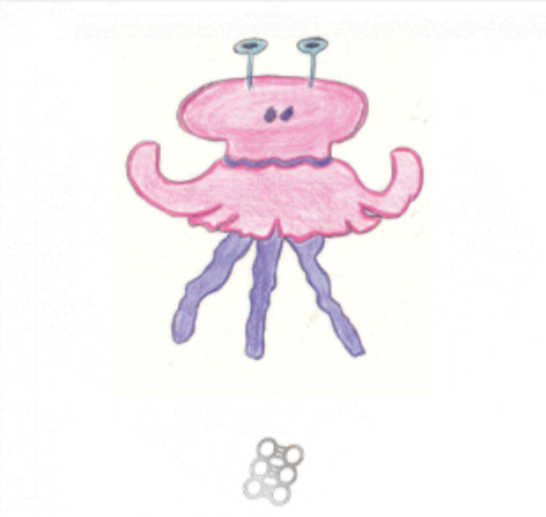


Which one of these is also a scred?

Bactras collect plastic rings.

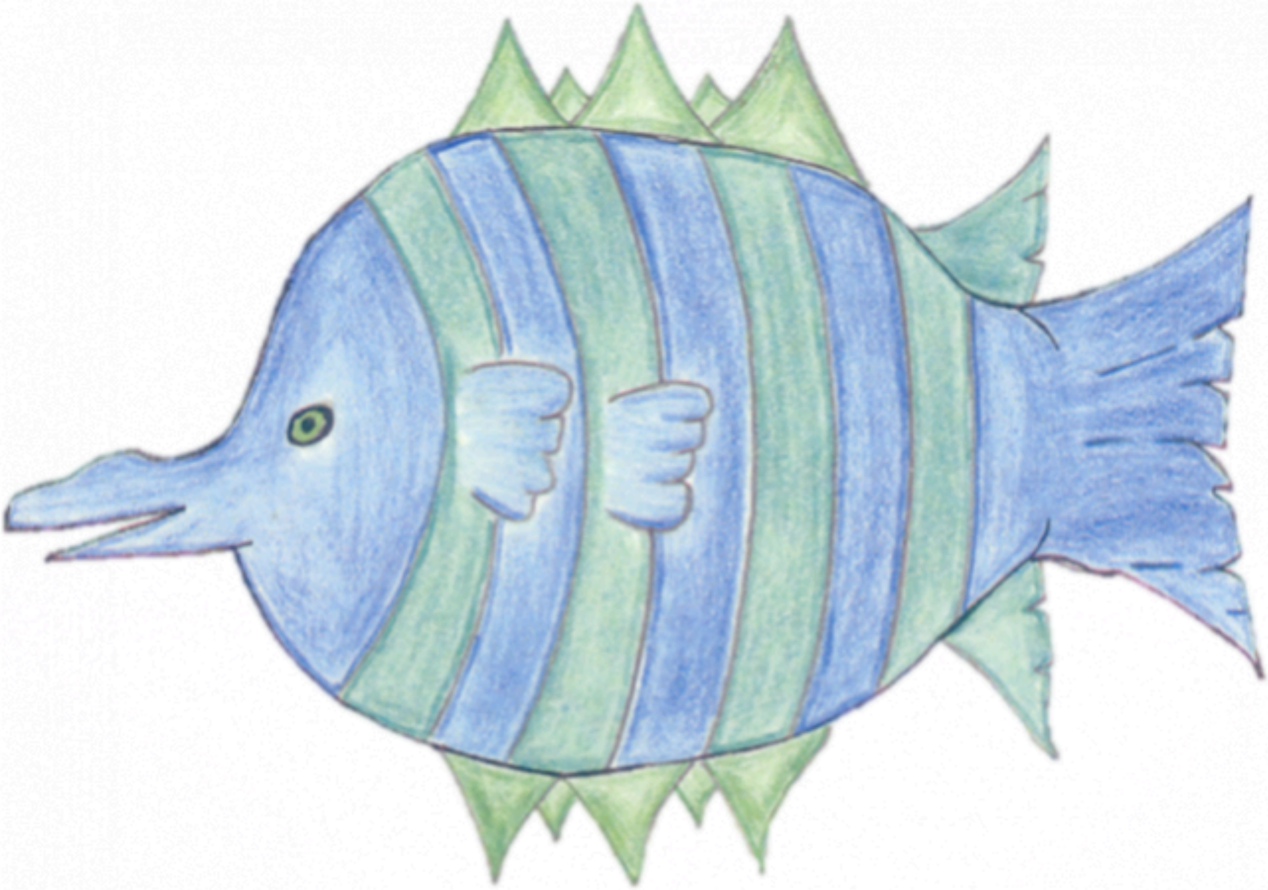


Which one of these is also a bactra?

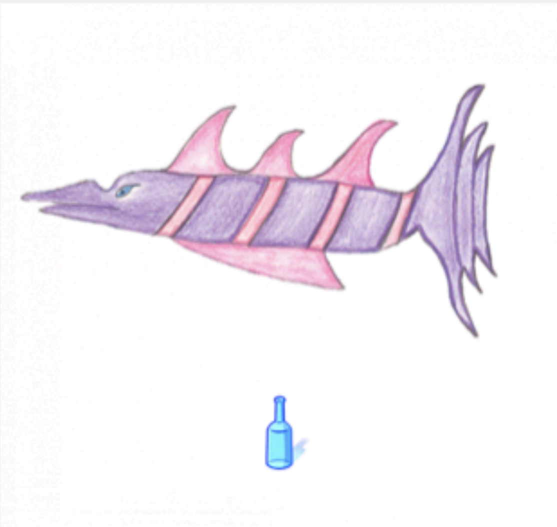


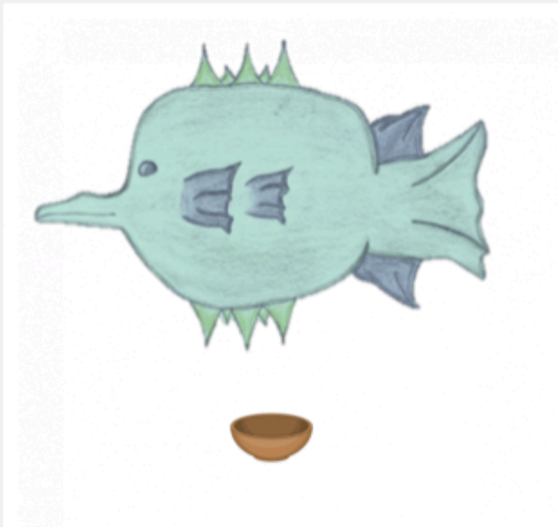
→

Bants have stripes.



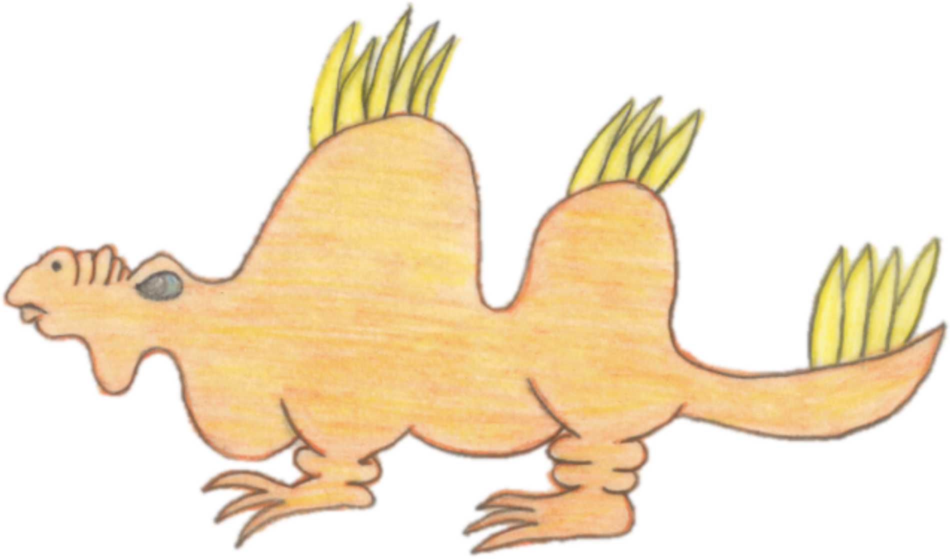
Which one of these is also a bant?

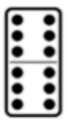




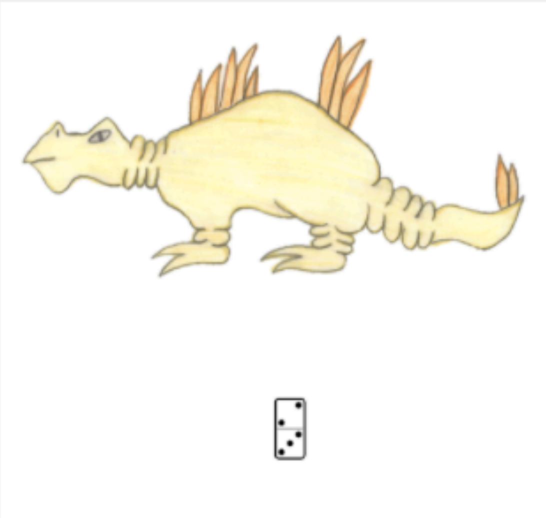
→

Febbits have spikes.



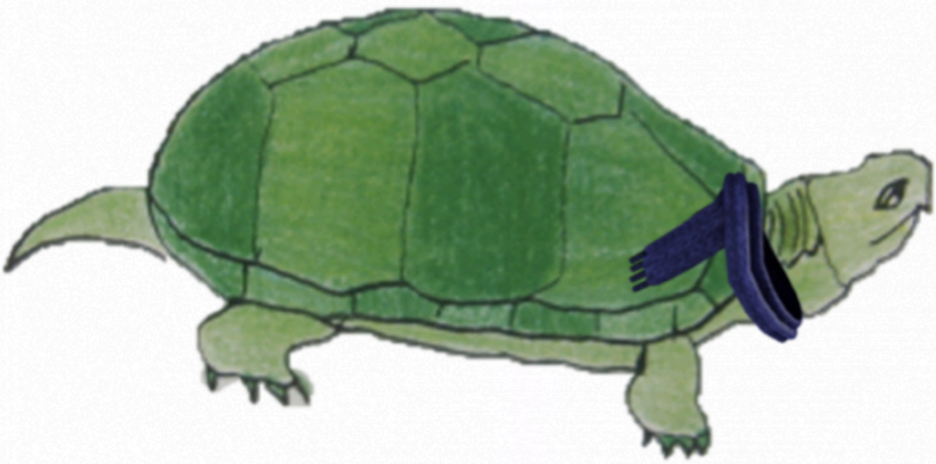


Which one of these is also a febbit?



→

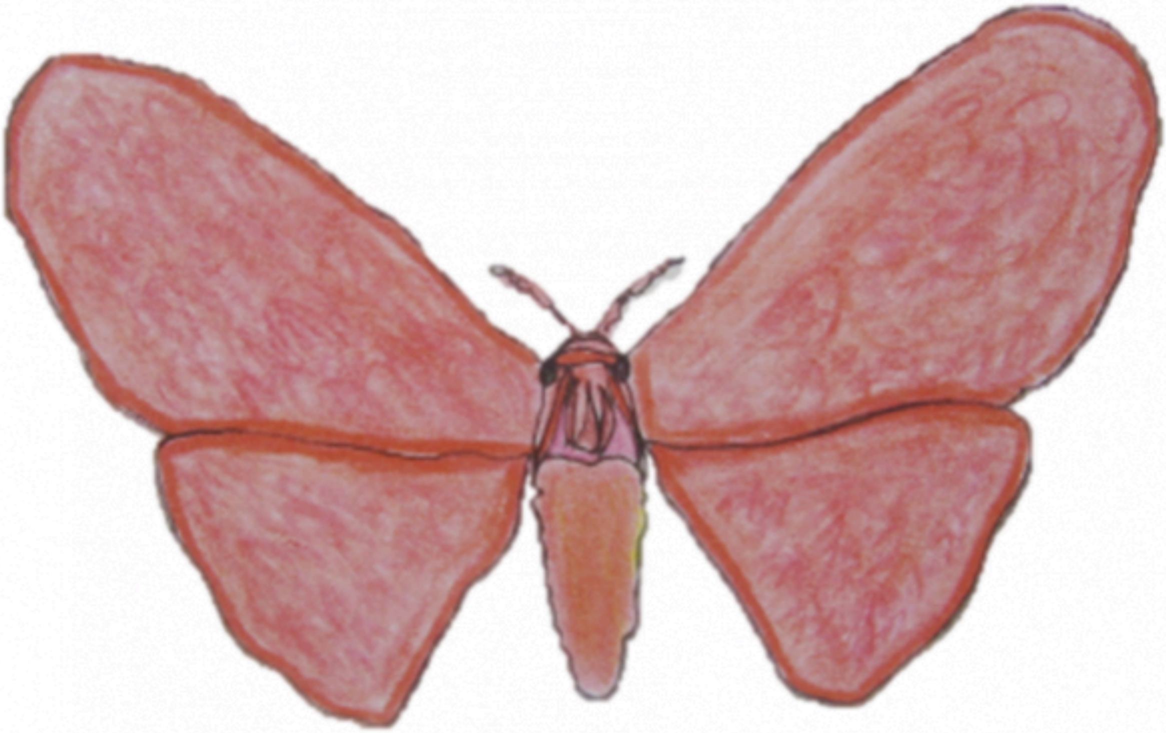
Plogs have green shells.



Which one of these is also a plog?

→

Vorzyds have red wings.



Which one of these is also a vorzyd?

→

Optional Demographics Section

The researchers would like to know a little bit about you. Please fill in the following part.
Note: This is an optional section.

Basic Information

What is your age?

Race/Ethnicity (Check all that apply)

- White
- Latino or Hispanic
- Black or African American
- Asian or Asian American
- American Indian or Alaska Native
- Middle Eastern or North African
- Native Hawaiian or Other Pacific Islander
- Not Listed

With which gender do you most identify?

Man

Woman

Transgender Man

Transgender Woman

Gender Non-Conforming

Not Listed

Prefer Not to Answer

What is your native language?

English

non-English

Did you live in an English-speaking environment from birth till (at least) age 13?

Yes

No

Did both of your parents speak English to you at home?

Yes

No

→

0% 100%

UNIVERSITY OF
MICHIGAN

OPTIONAL

Thank you for your participation. Before we get to the end of the experiment, we'd like to ask you one more question.
Is there anything you'd like to tell the researchers or anything else you noticed during the experiment?

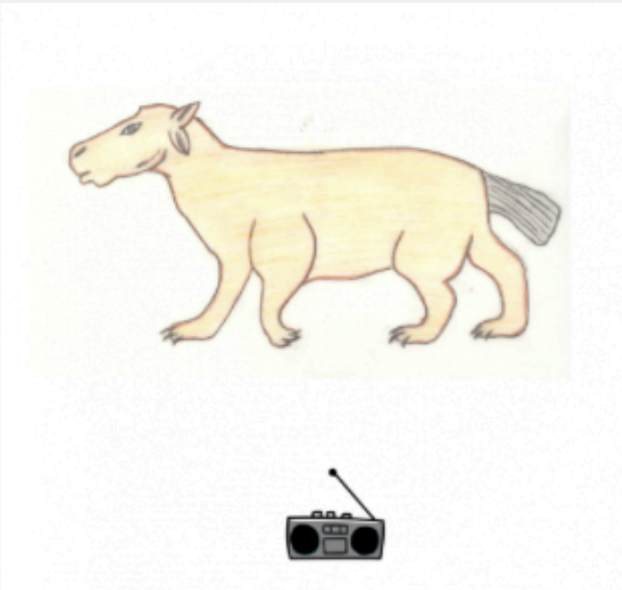
→

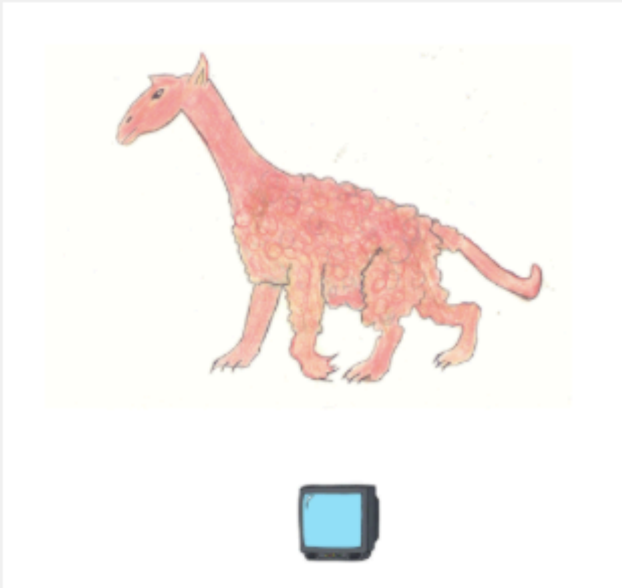
C.5.1 Sample stimuli for IS, DS, and control condition

Below is a sample of two trials each for the *kevtas* studies in the IS, DS, and 'this' subject conditions.

The kevta watches TV.

Which of these is also a kevta?





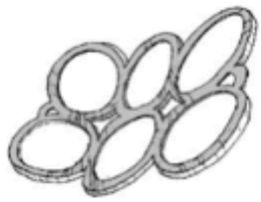
→

The fep wears a collar.

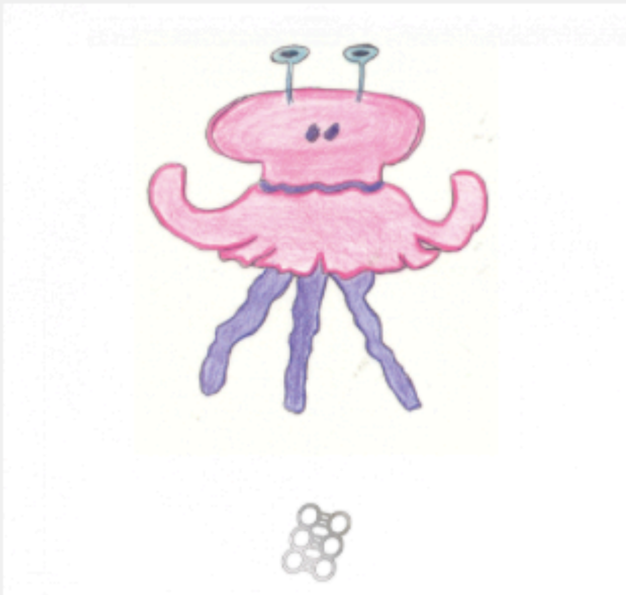
Which of these is also a fep?

→

This bactra has three legs.



Which one of these is also a bactra?



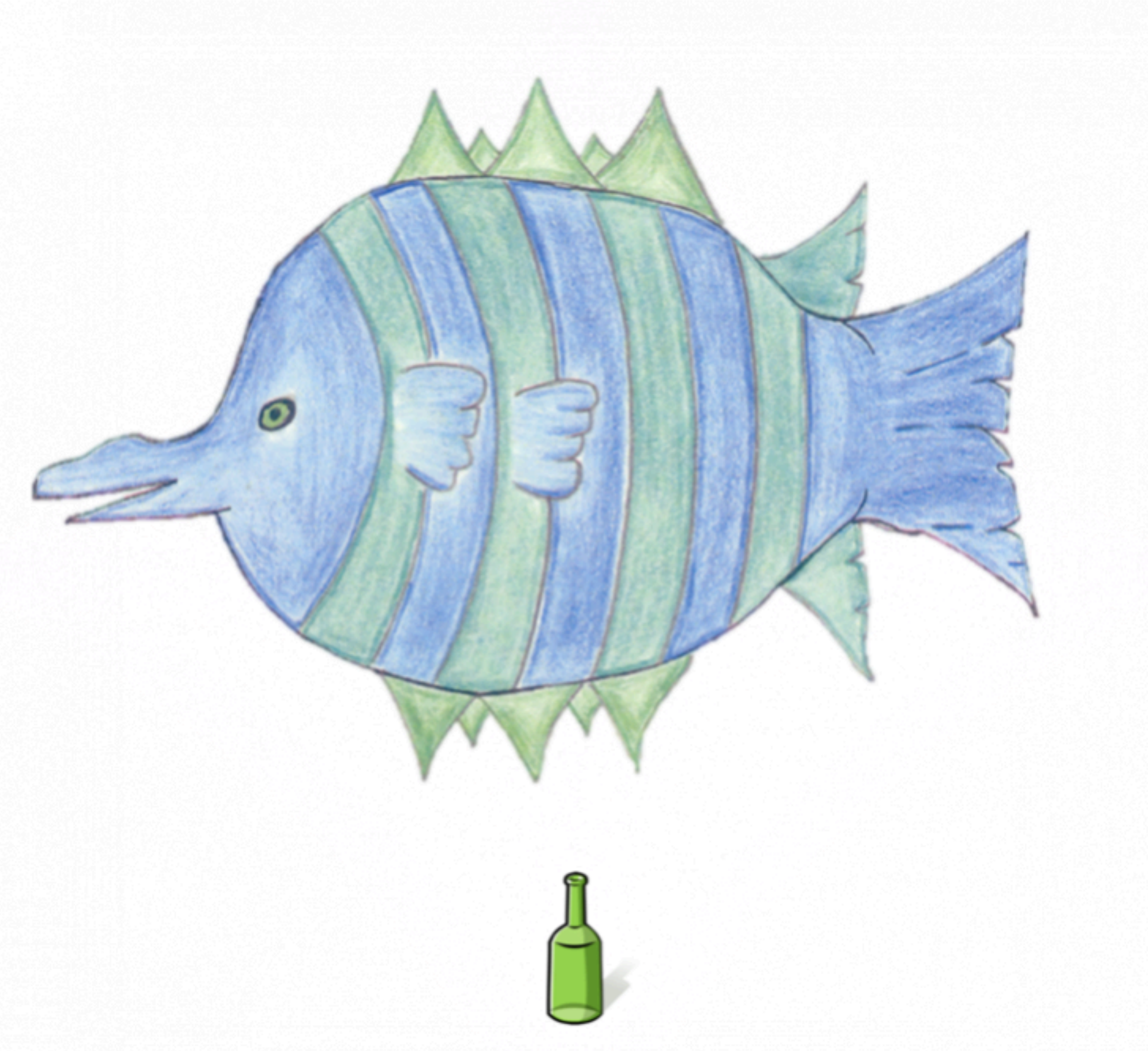
→

This fep is spotted.

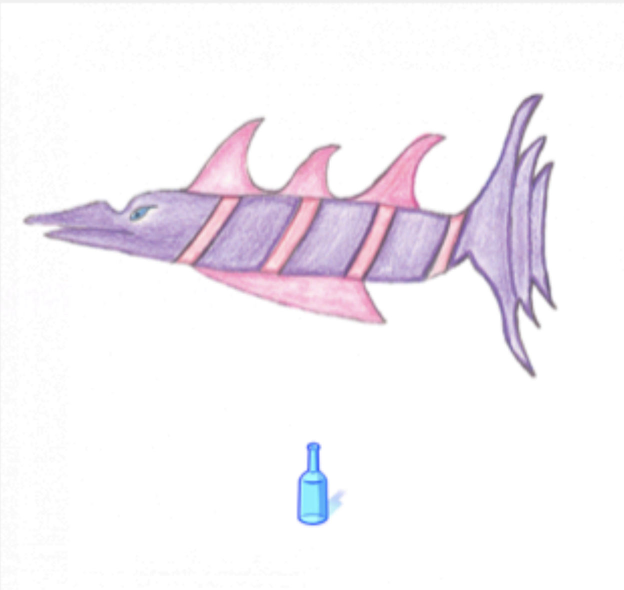
Which of these is also a fep?

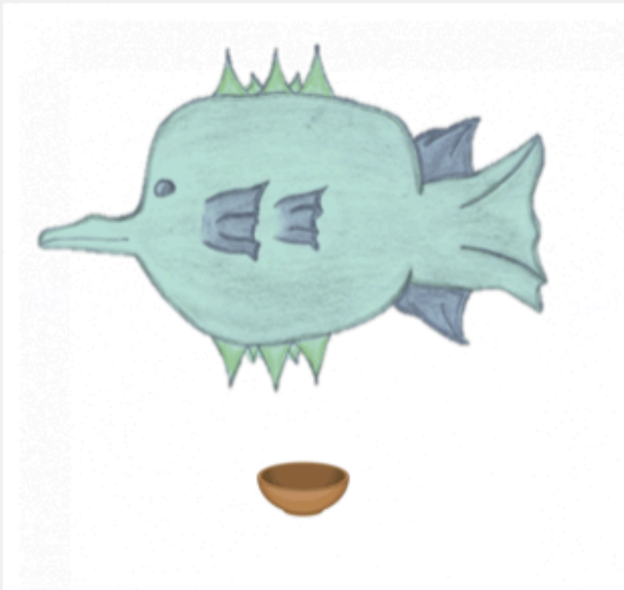
→

A bant has stripes.



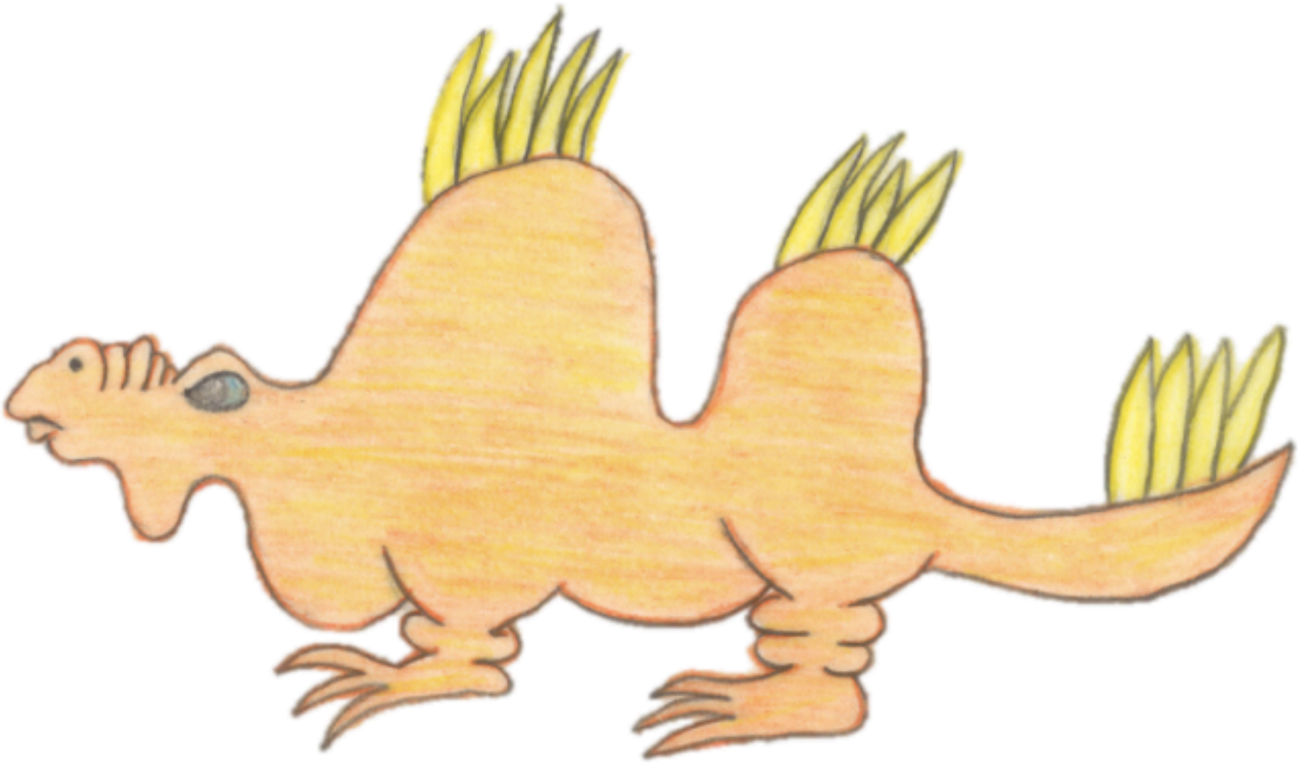
Which one of these is also a bant?

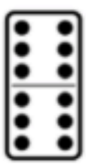




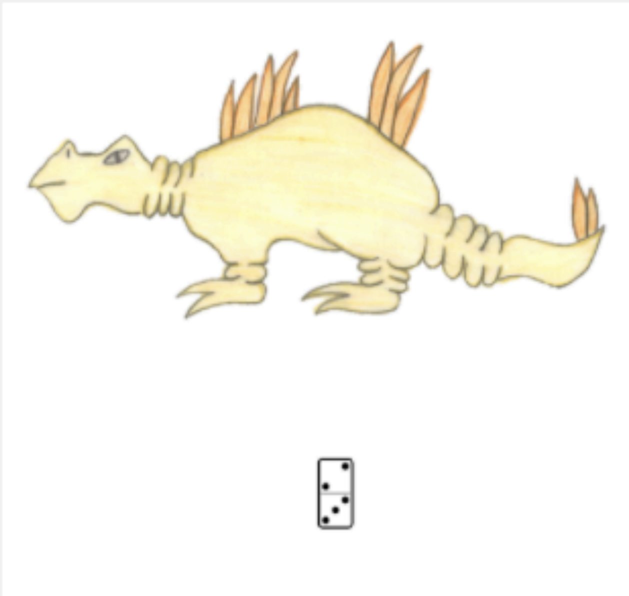
→

A febbit has spikes.





Which one of these is also a febbit?



→

D
Stimuli for Zarpies experiments

Contents

D.1 Participant Information Sheet

Participants received the following Information Sheet via email as part of the invitation to participate and in person before the start of the experiment in the Language & Brain Lab.

Faculty of Linguistics, Philology & Phonetics

Prof Matthew Husband
Email: matthew.husband@ling-phil.ox.ac.uk
Kim Fuellenbach
DPhil student
Clarendon Institute, Walton Street
Oxford, OX1 2HG
Email: kim.fuellenbach@ling-phil.ox.ac.uk

Thinking about new animals

PARTICIPANT INFORMATION SHEET

R55398/RE001

You are invited to take part in a research study investigating how people think about new animals and their particular properties. Please read this information to understand why the research is being done and what it would involve for you. If anything is unclear, or if you would like more information, please ask the researcher.

What is the purpose of the study?

The study investigates how people learn and think about new types of animals. In order to gain a better understanding of how the human conceptual system approaches this task, we provide information on a previously unknown type of animal. Through asking various questions about this new animal afterwards, we hope to better follow both the learning process and how people think about it.

Why have I been invited to take part?

You have been invited because you are a native speaker of English with no language impairments (dyslexia, aphasia), between 18 and 50 years of age.

Do I have to take part?

No. You may decide to stop being a part of the research study at any time without explanation and without penalty, by advising the researcher of this decision. You have the right to ask that any data you have supplied to that point be withdrawn or destroyed. You will still receive credit for your contribution.
You have the right to have your questions about the procedures answered (unless answering these questions would interfere with the study's outcome).

What will happen if I decide to take part?

You will be asked to read a picture book that describes novel animals and their properties in the Language and Brain Laboratory, either on your own or in a small group. You will then be asked to participate in several tasks related to the information presented in the picture book (e.g. comprehension questions) and unrelated tasks (e.g. solving basic math problems). Finally, you will complete a questionnaire about some beliefs you hold about the world. At the end of your visit, we can tell you more details about the specific purpose and goal of the study.

Are there any potential risks in taking part?
The experiment simply involves reading from a printed text and answering questions in a pen-and-paper manner. No specialised equipment or recording devices are used, and there are no additional risks in taking part in this research.

Participation in the experiment is voluntary, and you have the right to decline to participate or to withdraw at any time without penalty.

Will my taking part in the study be kept confidential?
Yes, the collected data will be analysed anonymously (using your participant number). No data will be linked to your personal information (e.g. email address).

Will I be reimbursed for taking part?
You will receive 3 credit points in return for your participation if you are participating through the Research Participation Scheme. If you are not participating through the PRS, you will be reimbursed £7.50 for your time.

What will happen to my data and the results of this study?
The data will be stored on an encrypted and password-protected computer. All research data and records will be stored for a minimum retention period of three years after publication or public release of the work of the research.

Will the research be published?
The results of this work will be used for presentation at conferences, publication in scientific journals and as part of the researcher's thesis. These results are always based on collected data from a large group of participants, such that data from individual participants are never identifiable.

Who is organising and funding the study?
The research is conducted by Kim Fuellenbach, D.Phil. candidate in General Linguistics and member of the Language and Brain Laboratory, part of the Faculty of Linguistics, Philology and Phonetics. This project is supervised by Prof. E. Matthew Husband.

Who has reviewed the study?
This study has been reviewed and approved by the Central University Research Ethics Committee (CUREC).

Participation in future research:
If you are willing to take part in future studies, please agree to this by ticking the optional box on the consent form.

Further information and contact details:
Please contact Kim Fuellenbach via kim.fuellenbach@ling-phil.ox.ac.uk.

Concerns or complaints
If you have a concern about any aspect of this study, please speak to the relevant researcher or their supervisor Prof Matthew Husband (matthew.husband@ling-phil.ox.ac.uk), who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the relevant chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter in a reasonably expeditious manner:

Chair, Social Sciences & Humanities Interdivisional Research Ethics Committee;
Email: ethics@socsci.ox.ac.uk;
Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD.

For further information and contact details, please contact:
Kim Fuellenbach
Clarendon Institute, Walton Street
Oxford, OX1 2HG
Email: kim.fuellenbach@ling-phil.ox.ac.uk

Thank you for reading this information.

D.2 Consent form

Participants signed the following consent form prior to beginning the study.

CONSENT FORM

Thinking about new animals

If you agree, please initial each box

		Please initial box
1	I confirm that I have read and understand the information sheet for this study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
2	I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without any adverse consequences or academic penalty.	
3	I understand that research data collected during the study may be looked at by designated individuals from the University of Oxford where it is relevant to my taking part in this study. I give permission for these individuals to access my data.	
4	I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee.	
5	I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project.	
6	I understand how this research will be written up and published.	

7	I understand how to raise a concern or make a complaint.	☐
8	I agree to take part in the above study.	☐
optional	I agree for my personal data to be kept in a secure database for the purpose of contacting me about future studies.	☐

Name of Participant: ____________________________

Date: ____________________________

Signature: ____________________________

Name of person taking Consent: ____________________________

Date: ____________________________

Signature: ____________________________

**1 copy for participant; 1 copy for researcher site file*

D.3 Stimuli

The following pages show the zarpies picture book in the Indefinite Singular wording condition. Subsection D.3.1 will show sample pages from the other wording conditions.

All About Zarpies!

Illustrated by
Dave Kush

Look at this zarpie!

A zarpie loves to eat tulips.

Look at this zarpie!

A zarpie has stripes on the bottom of its feet.

Look at this zarpie!

A zarpie can bounce a ball on its back.

Look at this zarpie!

A zarpie likes to sing.

Look at this zarpie!

A zarpie hides behind fences.

Look at this zarpie!

A zarpie wiggles its horns when it is happy.

Look at this zarpie!

A zarpie has spots on its tongue.

Look at this zarpie!

A zarpie hops over puddles.

Look at this zarpie!

A zarpie hates walking in snow.

Look at this zarpie!

A zarpie has a star behind its horns.

Look at this zarpie!

A zarpie can flip in the air.

Look at this zarpie!

A zarpie is scared of ladybugs.

Look at this zarpie!

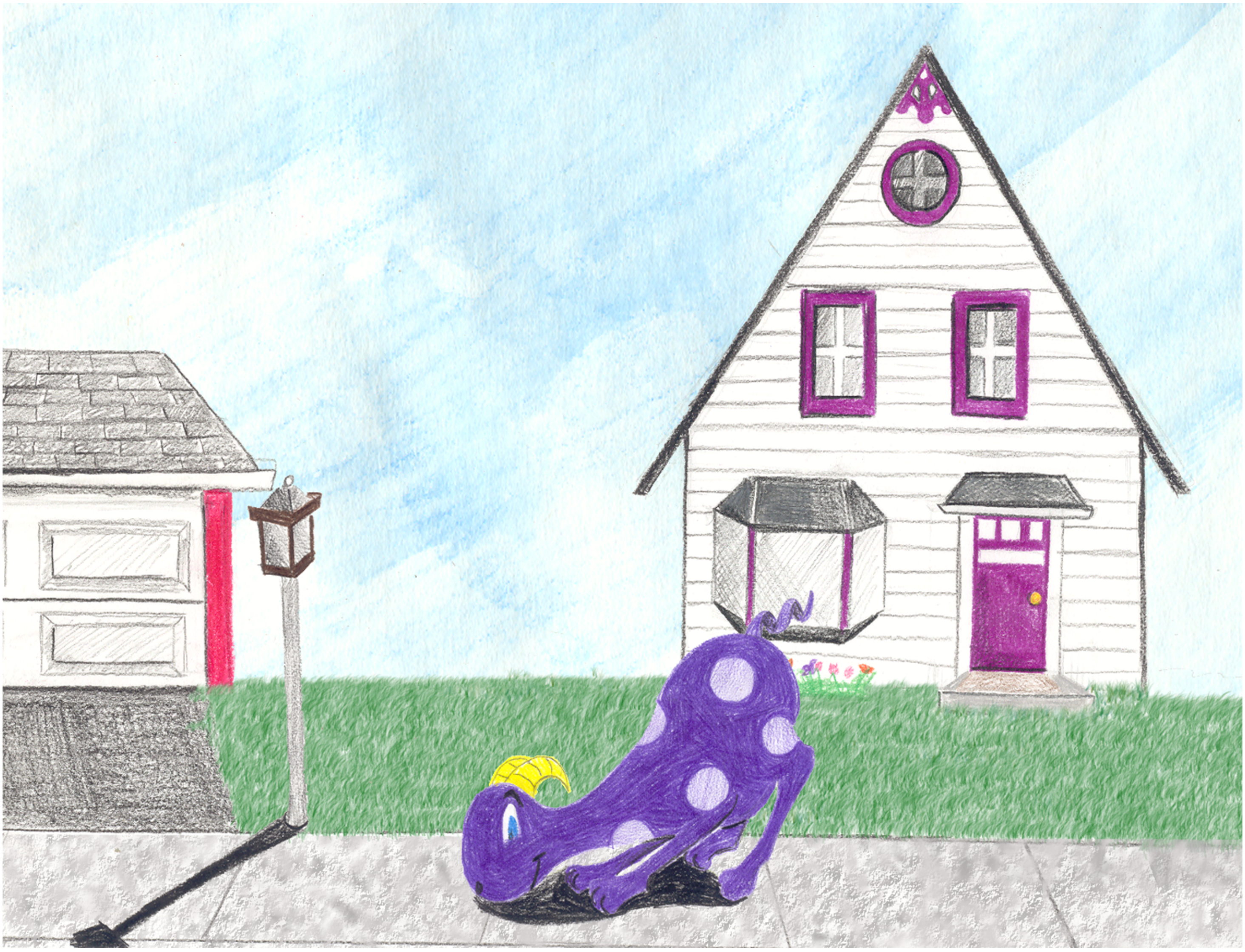
A baby zarpie has orange fur.

Look at this zarpie!

A zarpie stands on its back legs when it is surprised.

Look at this zarpie!

A zarpie chases shadows.



Look at this zarpie!

A zarpie hates ice cream.


ICE
CREAM

Look at this zarpie!

A zarpie sleeps in tall trees.


zZZ

D.3.1 Further stimuli

To illustrate the zarpies picture in the BP, DS, and most-quantified wording conditions, the following section shows a sample page of each of the three other wording conditions that were used in the zarpies studies.

Look at this zarpie!

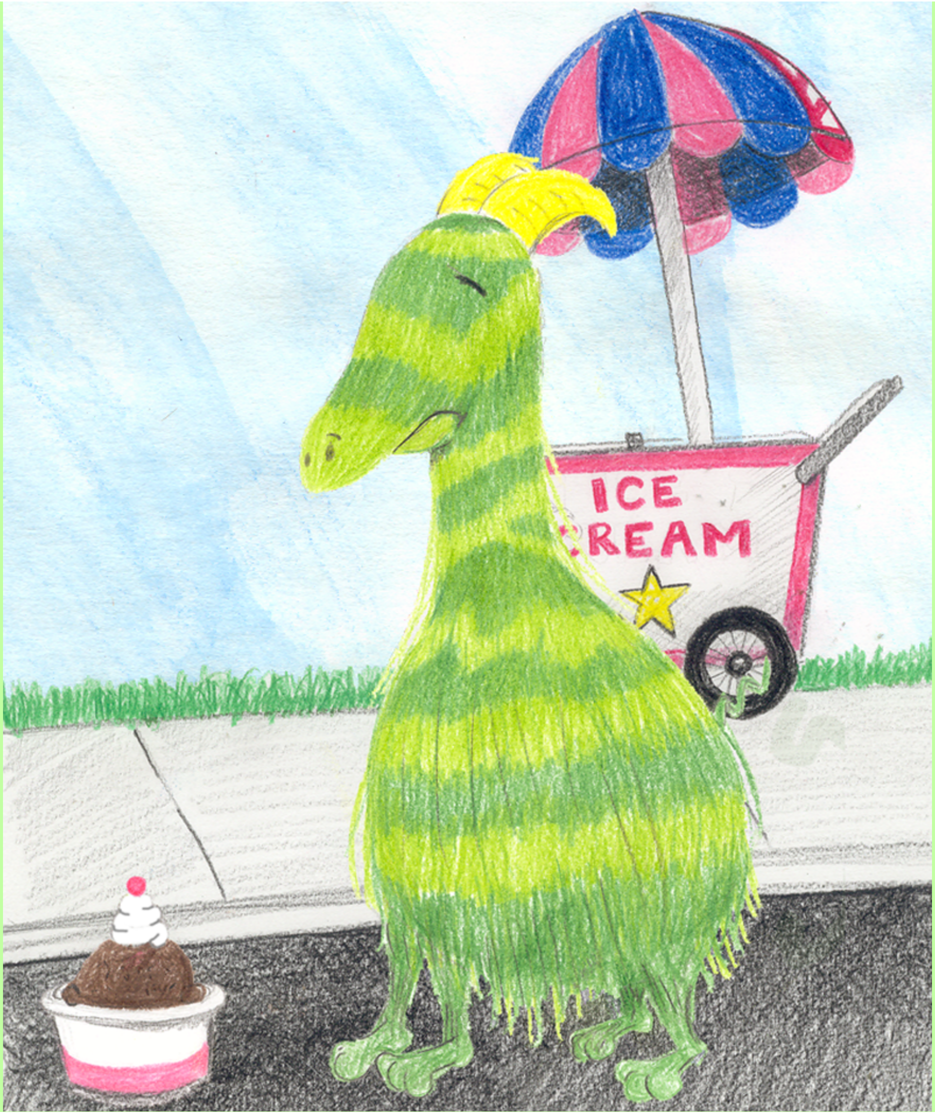
Zarpies can bounce a ball on their backs.

Look at this zarpie!

The zarpie wiggles its horns when it is happy.

Look at this zarpie!

Most zarpies hate ice cream.


ICE
CREAM

D.4 Test booklet

After looking through the picture book for five minutes, participants were asked to complete the following test booklet in their own time.

Participant #________ Date__________________ Form AG

Please answer the following questions.

1a. One of these pictures was in the book, the other was not.
Please circle the one that was in the book.

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date___________________ Form AG

1b. One of these pictures was in the book, the other was not.
Please circle the one that was in the book.

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

1c. One of these pictures was in the book, the other was not.
Please circle the one that was in the book.

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

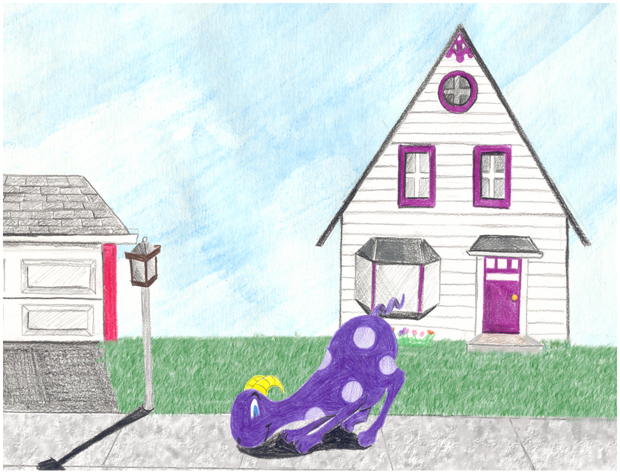
1d. One of these pictures was in the book, the other was not.
Please circle the one that was in the book.

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

Please answer the following questions to the best of your ability.

2a. Why is this chasing a shadow?



Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

2b. Why does this hate walking in snow?

____

____

____

____

____

____

Please continue to the next page. **DO NOT** skip ahead or turn back.

2c. Why is this hiding behind a fence?

____

____

____

____

____

____

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

2d. Why is this sleeping in a tall tree?

____

____

____

____

____

____

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

Please circle one answer choice for each question.

3a. Look at this zarpie.

This zarpie is scared of ladybugs. You know what else? This zarpie is 8 years old.

(1) Do you think it was scared of ladybugs when it was 4 years old? Yes No

(2) Do you think it will always be scared of ladybugs? Yes No

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

3b. Look at this zarpie.

This zarpie can flip in the air. You know what else? This zarpie is 8 years old.

(1) Do you think it could flip in the air when it was 4 years old? Yes No

(2) Do you think it will always be able to flip in the air? Yes No

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

3c. Look at this zarpie.

This zarpie hops over puddles. You know what else? This zarpie is 8 years old.

(1) Do you think it hopped over puddles when it was 4 years old? Yes No

(2) Do you think it will always hop over puddles? Yes No

Please continue to the next page. **DO NOT** skip ahead or turn back.

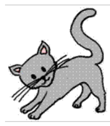
Participant #________ Date__________________ Form AG

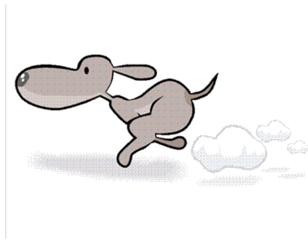
Please read the following stories and answer the questions.

4a.

Look at this zarpie.
She chases shadows.

(1) One day this zarpie had a baby.
That means that the baby came
out of her tummy.





Look at this dog.
She chases cats.

(2) But, right after the baby was
born, it went to live with this dog.
(3)The dog took care of the baby.
She played with the baby, fed the
baby, and loved the baby.
(4)The baby grew up with this
dog…

(5) …and never saw this zarpie again.

Now the baby is 6 years old.
Do you think that the baby:

…chases shadows, like this zarpie? **OR** …chases cats, like this dog?

☐ ☐

(Please check the appropriate box).

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #_______ Date________________ Form AG

4b.

Look at this zarpie.
She likes to sing.

Look at this dog.
She likes to bark.

(1) One day this zarpie had a baby. That means that the baby came out of her tummy.

(2) But, right after the baby was born, it went to live with this dog.
(3)The dog took care of the baby. She played with the baby, fed the baby, and loved the baby.
(4)The baby grew up with this dog…

(5) …and never saw this zarpie again.

Now the baby is 6 years old.
Do you think that the baby:

…likes to sing, like this zarpie? **OR** …likes to bark, like this dog?

☐ ☐

(Please check the appropriate box).

Please continue to the next page. **DO NOT** skip ahead or turn back.

4c.

Look at this zarpie.
She wiggles her horns
when she is happy.

Look at this dog.
She wags her tail
when she is happy.

(1) One day this zarpie had a baby.
That means that the baby came
out of her tummy.

(2) But, right after the baby was born, it went to live with this dog. (3)The dog took care of the baby. She played with the baby, fed the baby, and loved the baby. (4)The baby grew up with this dog…

(5) …and never saw this zarpie again.

Now the baby is 6 years old.
Do you think that the baby:

…wiggles its horns when it is happy, like this zarpie? ☐

OR

…wags its tail when it is happy, like this dog? ☐

(Please check the appropriate box).

Please continue to the next page. **DO NOT** skip ahead or turn back.

4d.

Look at this zarpie.
She loves to eat tulips.

Look at this dog.
She loves to chew bones.

(1) One day this zarpie had a baby.
That means that the baby came
out of her tummy.

(2) But, right after the baby was born, it went to live with this dog.
(3)The dog took care of the baby. She played with the baby, fed the baby, and loved the baby.
(4)The baby grew up with this dog…

(5) …and never saw this zarpie again.

Now the baby is 6 years old.
Do you think that the baby:

…loves to eat tulips, like this zarpie? **OR** …loves to chew bones, like this dog?

☐ ☐

(Please check the appropriate box).

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

Please circle one answer choice for each question.

5a. Remember this zarpie?

Can you find another zarpie? Please circle one answer choice.

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

5b. Remember this zarpie?

Can you find another zarpie? Please circle one answer choice.

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

5c. Remember this zarpie?

Can you find another zarpie? Please circle one answer choice.

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

5d. Remember this zarpie?

Can you find another zarpie? Please circle one answer choice.

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

Please answer the following questions.

6a. Remember this zarpie?

This zarpie can bounce a ball on its back.

(1) Look at this zarpie.

Do you think that this zarpie can bounce a ball on its back? Yes No

(2) Look at this zarpie.

Do you think that this zarpie can bounce a ball on its back? Yes No

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

6b. Remember this zarpie?

This zarpie stands on its back legs when it is surprised.

(1) Look at this zarpie.

Do you think that this zarpie stands on its back legs when it is surprised? Yes No

(2) Look at this zarpie.

Do you think that this zarpie stands on its back legs when it is surprised? Yes No

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

6c. Remember this zarpie?

This zarpie hates ice cream.

(1) Look at this zarpie.

Do you think that this zarpie hates ice cream? Yes No

(2) Look at this zarpie.

Do you think that this zarpie hates ice cream? Yes No

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

Please answer the following questions.

7a. Look at this zarpie.

This zarpie dances in circles.

(1) Look at this zarpie.

Do you think that this zarpie dances in circles? Yes No

(2) Look at this zarpie.

Do you think that this zarpie dances in circles? Yes No

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

7b. Look at this zarpie.

This zarpie makes a buzzing sound when it is angry.

(1) Look at this zarpie.

Do you think that this zarpie makes a buzzing sound when it is angry? Yes No

(2) Look at this zarpie.

Do you think that this zarpie makes a buzzing sound when it is angry? Yes No

Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

7c. Look at this zarpie.

This zarpie likes to play with Frisbees.

(1) Look at this zarpie.

Do you think that this zarpie likes to play with Frisbees? Yes No

(2) Look at this zarpie.

Do you think that this zarpie likes to play with Frisbees? Yes No

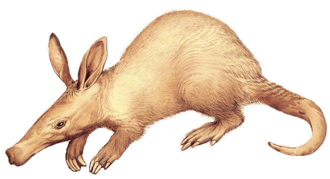
Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

Please circle one answer choice for each question.

8a. See this zarpie? This zarpie is a blicket.

Can you find another blicket? Please circle one answer choice.

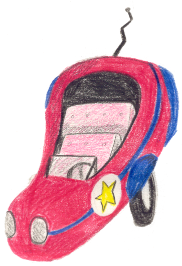


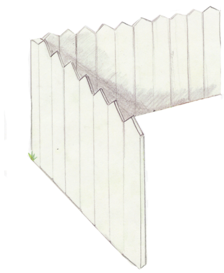
Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

8b. See this zarpie? This zarpie is a meeber.

Can you find another meeber? Please circle one answer choice.



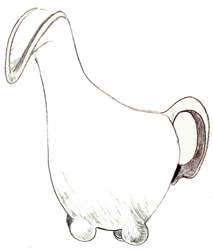


Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

8c. See this zarpie? This zarpie is a fendle.

Can you find another fendle? Please circle one answer choice.

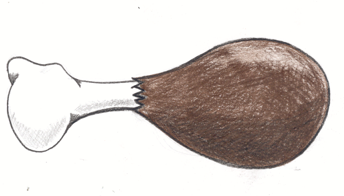


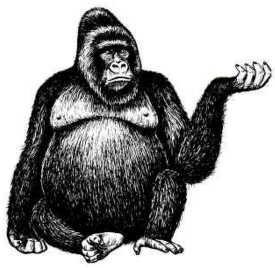
Please continue to the next page. **DO NOT** skip ahead or turn back.

Participant #________ Date__________________ Form AG

8d. See this zarpie? This zarpie is a toma.

Can you find another toma? Please circle one answer choice.





Please continue to the next page. **DO NOT** skip ahead or turn back.

D.5 Essentialism questionnaire

At the end of the study, participants were asked to fill out this essentialism questionnaire. This questionnaire was adapted from Haslam et al. (2000) and Gelman et al. (2007).

Today's date ___/___/___ Participant #___________________

Your gender: ________________________ Form A

1.

B. Are some people born with a predisposition to be <u>artistic</u>?
1 2 3 4 5 6
definitely no definitely yes

P. Are some people born with a predisposition to be <u>pessimistic</u>?
1 2 3 4 5 6
definitely no definitely yes

R. Are some people born with a predisposition to have <u>a good sense of direction</u>?
1 2 3 4 5 6
definitely no definitely yes

D. Are some people born with a predisposition to be <u>cautious</u>?
1 2 3 4 5 6
definitely no definitely yes

K. Are some people born with a predisposition to be <u>messy</u>?
1 2 3 4 5 6
definitely no definitely yes

A. Are some people born with a predisposition to be <u>anxious</u>?
1 2 3 4 5 6
definitely no definitely yes

G. Are some people born with a predisposition to be <u>humorous</u>?
1 2 3 4 5 6
definitely no definitely yes

Q. Are some people born with a predisposition to be <u>responsible</u>?
1 2 3 4 5 6
definitely no definitely yes

J. Are some people born with a predisposition to be <u>mathematically talented</u>?
1 2 3 4 5 6
definitely no definitely yes

C. Are some people born with a predisposition to be <u>athletic</u>?
1 2 3 4 5 6
definitely no definitely yes

F. Are some people born with a predisposition to be <u>goal oriented</u>?
1 2 3 4 5 6
definitely no definitely yes

N. Are some people born with a predisposition to be **nurturing**?

1	2	3	4	5	6
definitely no					definitely yes

I. Are some people born with a predisposition to be **masculine**?

1	2	3	4	5	6
definitely no					definitely yes

O. Are some people born with a predisposition to be **outgoing**?

1	2	3	4	5	6
definitely no					definitely yes

T. Are some people born with a predisposition to be **stubborn**?

1	2	3	4	5	6
definitely no					definitely yes

E. Are some people born with a predisposition to be **feminine**?

1	2	3	4	5	6
definitely no					definitely yes

H. Are some people born with a predisposition to be **intelligent**?

1	2	3	4	5	6
definitely no					definitely yes

M. Are some people born with a predisposition to be **neat**?

1	2	3	4	5	6
definitely no					definitely yes

L. Are some people born with a predisposition to be **musically talented**?

1	2	3	4	5	6
definitely no					definitely yes

S. Are some people born with a predisposition to be **shy**?

1	2	3	4	5	6
definitely no					definitely yes

2.

J. Will a baby who is adopted at birth grow up to be as **mathematically talented** as the birth parents, or as **mathematically talented** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

R. Will a baby who is adopted at birth grow up to have as **good of a sense of direction** as the birth parents, or as **good of a sense of direction** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

S. Will a baby who is adopted at birth grow up to be as **shy** as the birth parents, or as **shy** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

C. Will a baby who is adopted at birth grow up to be as **athletic** as the birth parents, or as **athletic** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

M. Will a baby who is adopted at birth grow up to be as **neat** as the birth parents, or as **neat** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

D. Will a baby who is adopted at birth grow up to be as **cautious** as the birth parents, or as **cautious** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

P. Will a baby who is adopted at birth grow up to be as **pessimistic** as the birth parents, or as **pessimistic** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

G. Will a baby who is adopted at birth grow up to be as **humorous** as the birth parents, or as **humorous** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

Q. Will a baby who is adopted at birth grow up to be as **responsible** as the birth parents, or as **responsible** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

I. Will a baby who is adopted at birth grow up to be as **masculine** as the birth parents, or as **masculine** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

H. Will a baby who is adopted at birth grow up to be as **intelligent** as the birth parents, or as **intelligent** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

E. Will a baby who is adopted at birth grow up to be as **feminine** as the birth parents, or as **feminine** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

O. Will a baby who is adopted at birth grow up to be as **outgoing** as the birth parents, or as **outgoing** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

F. Will a baby who is adopted at birth grow up to be as **goal oriented** as the birth parents, or as **goal oriented** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

B. Will a baby who is adopted at birth grow up to be as **artistic** as the birth parents, or as **artistic** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

A. Will a baby who is adopted at birth grow up to be as **anxious** as the birth parents, or as **anxious** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

T. Will a baby who is adopted at birth grow up to be as **stubborn** as the birth parents, or as **stubborn** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

L. Will a baby who is adopted at birth grow up to be as **musically talented** as the birth parents, or as **musically talented** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

K. Will a baby who is adopted at birth grow up to be as **messy** as the birth parents, or as **messy** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

N. Will a baby who is adopted at birth grow up to be as **nurturing** as the birth parents, or as **nurturing** as the adoptive parents?

1	2	3	4	5	6
adoptive parents					birth parents

M. Do people who are <u>neat</u> at age 20 tend to be <u>neat</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

J. Do people who are <u>mathematically talented</u> at age 20 tend to be <u>mathematically talented</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

E. Do people who are <u>feminine</u> at age 20 tend to be <u>feminine</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

R. Do people who have <u>a good sense of direction</u> at age 20 tend to have <u>a good sense of direction</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

Q. Do people who are <u>responsible</u> at age 20 tend to be <u>responsible</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

F. Do people who are <u>goal oriented</u> at age 20 tend to be <u>goal oriented</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

O. Do people who are <u>outgoing</u> at age 20 tend to be <u>outgoing</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

K. Do people who are <u>messy</u> at age 20 tend to be <u>messy</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

D. Do people who are <u>cautious</u> at age 20 tend to be <u>cautious</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

C. Do people who are <u>athletic</u> at age 20 tend to be <u>athletic</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

H. Do people who are <u>intelligent</u> at age 20 tend to be <u>intelligent</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

A. Do people who are <u>anxious</u> at age 20 tend to be <u>anxious</u> at age 40?

1	2	3	4	5	6
definitely no					definitely yes

I. Do people who are **masculine** at age 20 tend to be **masculine** at age 40?

1	2	3	4	5	6
definitely no					definitely yes

S. Do people who are **shy** at age 20 tend to be **shy** at age 40?

1	2	3	4	5	6
definitely no					definitely yes

N. Do people who are **nurturing** at age 20 tend to be **nurturing** at age 40?

1	2	3	4	5	6
definitely no					definitely yes

L. Do people who are **musically talented** at age 20 tend to be **musically talented** at age 40?

1	2	3	4	5	6
definitely no					definitely yes

T. Do people who are **stubborn** at age 20 tend to be **stubborn** at age 40?

1	2	3	4	5	6
definitely no					definitely yes

B. Do people who are **artistic** at age 20 tend to be **artistic** at age 40?

1	2	3	4	5	6
definitely no					definitely yes

G. Do people who are **humorous** at age 20 tend to be **humorous** at age 40?

1	2	3	4	5	6
definitely no					definitely yes

P. Do people who are **pessimistic** at age 20 tend to be **pessimistic** at age 40?

1	2	3	4	5	6
definitely no					definitely yes

6.

F. How much does the environment a person grows up in, affect whether or not he or she will become **goal oriented**?

1	2	3	4	5	6
not at all					a lot

N. How much does the environment a person grows up in, affect whether or not he or she will become **nurturing**?

1	2	3	4	5	6
not at all					a lot

Q. How much does the environment a person grows up in, affect whether or not he or she will become **responsible**?

1	2	3	4	5	6
not at all					a lot

P. How much does the environment a person grows up in, affect whether or not he or she will become **pessimistic**?

1	2	3	4	5	6
not at all					a lot

B. How much does the environment a person grows up in, affect whether or not he or she will become **artistic**?

1	2	3	4	5	6
not at all					a lot

J. How much does the environment a person grows up in, affect whether or not he or she will become **mathematically talented**?

1	2	3	4	5	6
not at all					a lot

D. How much does the environment a person grows up in, affect whether or not he or she will become **cautious**?

1	2	3	4	5	6
not at all					a lot

M. How much does the environment a person grows up in, affect whether or not he or she will become **neat**?

1	2	3	4	5	6
not at all					a lot

H. How much does the environment a person grows up in, affect whether or not he or she will become **intelligent**?

1	2	3	4	5	6
not at all					a lot

I. How much does the environment a person grows up in, affect whether or not he or she will become **masculine**?

1	2	3	4	5	6
not at all					a lot

T. How much does the environment a person grows up in, affect whether or not he or she will become

stubborn?

1 2 3 4 5 6
not at all a lot

C. How much does the environment a person grows up in, affect whether or not he or she will become **athletic**?

1 2 3 4 5 6
not at all a lot

S. How much does the environment a person grows up in, affect whether or not he or she will become **shy**?

1 2 3 4 5 6
not at all a lot

A. How much does the environment a person grows up in, affect whether or not he or she will become **anxious**?

1 2 3 4 5 6
not at all a lot

G. How much does the environment a person grows up in, affect whether or not he or she will become **humorous**?

1 2 3 4 5 6
not at all a lot

R. How much does the environment a person grows up in, affect whether or not he or she will have **a good sense of direction**?

1 2 3 4 5 6
not at all a lot

L. How much does the environment a person grows up in, affect whether or not he or she will become **musically talented**?

1 2 3 4 5 6
not at all a lot

K. How much does the environment a person grows up in, affect whether or not he or she will become **messy**?

1 2 3 4 5 6
not at all a lot

O. How much does the environment a person grows up in, affect whether or not he or she will become **outgoing**?

1 2 3 4 5 6
not at all a lot

E. How much does the environment a person grows up in, affect whether or not he or she will become **feminine**?

1 2 3 4 5 6
not at all a lot

E
Stimuli for Reverse Kevtas experiments

Contents

E.1 Participant Information Sheet and Consent Form

The consent form was adapted from the *kevtas* studies, in line with the demands by the university's ethics committee (CUREC). This form was approved under the same ethics application (CUREC R55398-RE001 through R55398-RE004).

Consent to Participate in Psycholinguistics Study

This study is being conducted by Matthew Husband and Kim Fuellenbach, researchers at the University of Oxford.

Invitation to Participate in Our Study

You are invited to take part in a research study investigating how people think about new animals and their particular properties. Please read this information to understand why the research is being done and what it would involve for you. If anything is unclear, or if you would like more information, please ask the researcher.

You are invited because you are a native speaker of English. You are being asked to take part in a research study which explores how we think about the world.

Study Procedure

We invite you to fill out this questionnaire. The study will take around 8 minutes to complete, and involves taking a survey on a computer. You will answer some questions about animals.

Risks of Participation
The risks of participation in this online study are those associated with basic computer tasks, boredom, fatigue, mild stress or breach of confidentiality. If you are uncomfortable for any reason and wish to end participation, you are free to stop at any time. To stop, simply close the browser window.

Benefits of Participation
There is no direct benefit to you that you can expect to receive as a result of participating in this study. However, we ultimately hope that our findings can add in helpful ways to what we know about child development.

Confidentiality of Records
We hope to publish the results of this study but will not include information that identifies you. Your Mechanical Turk Worker ID will be stored separately from the data we collect from you. Please be aware that your MTurk Worker ID can be potentially linked to information about you on your Amazon public profile page, depending on settings you have for you Amazon profile. We will not be accessing any personally identifying information about you that you may have put on your Amazon public profile page.

Voluntary Participation
Your participation in this study is voluntary. You are to stop participating at any time by closing the browser window. In the event that you close your browser during the study, the information you have already provided will not be analyzed.

Do you consent to participate in our study?

Yes, I acknowledge that I am at least 18, I am a native speaker of English and consent to participate in this research study.

E.2 Demographics and essentialism questionnaires

The demographics and essentialism questionnaires were the same as the ones for the *kevtas* studies. As in those studies, completion thereof was optional.

E.3 Stimuli

Below is a sample trial of the reverse kevtas study, including the welcome message and instructions that participants received after providing consent and three sample trials. This study was hosted on Qualtrics and was designed for adult participants. Note that this is only one of the possible randomised trials.

Welcome!

In this task, you will be seeing pictures of some new animals. They are different from the animals we know, so you will learn the name for each of them. For each animal, you will be given its name and see a picture of it along with a smaller image of something in its environment.

Then you will see two more pictures. One has a circle around it and is the same kind of animal as the animal in the first picture. The other one has no circle and is not the same kind of animal.

You will be asked to choose a sentence that best describes the first image so that someone else would know which of the other two animals was the same kind as the first one.

Please answer to the best of your ability.

When you're ready, please turn off any distractions, such as your cell phone, for the duration of the experiment! Please avoid using the mobile version of this site.

Please enter your Prolific ID here:

0% 100%

This picture is of a fep.

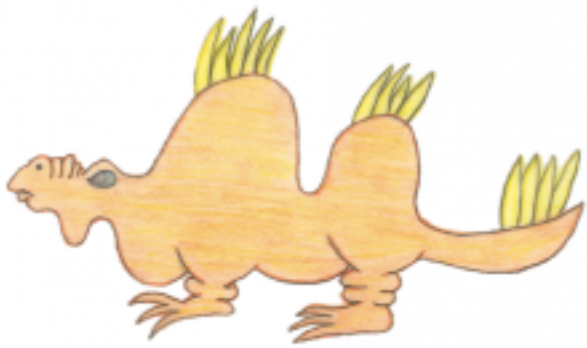
Here are two more pictures. The one with the circle is also a fep. The other one is not.

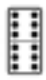
Which of the following four options would you use to describe the first picture to someone, so that they would also think the circled picture is a fep, and not the other one?

- The fep is spotted.
- A fep is spotted.
- This fep is spotted.
- None of these.

→

This picture is of a febbit.





Here are two more pictures. The one with the circle is also a febbit. The other one is not.

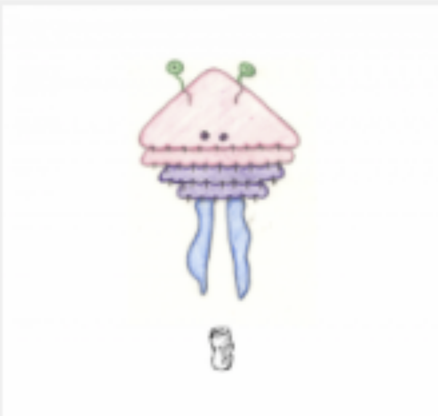
Which of the following four options would you use to describe the first picture to someone, so that they would also think the circled picture is a febbit, and not the other one?

- A febbit has spikes.
- The febbit has spikes.
- None of these.
- This febbit has spikes.

→

This is a picture of a bactra.

Here are two more pictures. The one with the circle is also a bactra.
The other one is not.



Which of the following four options would you use to describe the first picture to someone, so that they would also think the circled picture is a bactra, and not the other one?

- ◯ The bactra has three legs.
- ◯ This bactra has three legs.
- ◯ A bactra has three legs.
- ◯ None of these.

→

E.4 Further analyses

Table E.1: Parameter estimates of Experiment 15 (production), for log odds of IS/None, DS/None, This/None

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept [IS/None]	-0.381	0.118	10.37	0.0013*
Connection Type[PC]	0.296	0.118	6.29	0.0121*
Match type[prop]	2.317	0.118	384.19	<.0001*
Connection Type[PC] × Match type[prop]	0.328	0.118	7.68	0.0056*
Intercept [DS/None]	-0.997	0.136	54.03	<.0001*
Connection Type[PC]	-0.204	0.136	2.27	0.1319
Match type[prop]	2.091	0.136	237.53	<.0001*
Connection Type[PC] × Match type[prop]	0.220	0.136	2.63	0.1047
Intercept [This/None]	-0.853	0.116	54.13	<.0001*
Connection Type[PC]	0.23	0.116	3.92	0.0476*
Match type[prop]	0.997	0.116	74.01	<.0001*
Connection Type[PC] × Match type[prop]	-0.011	0.116	0.01	0.9233