



PROJECT MUSE®

Verbs of perception: A quantitative typological study

Elisabeth Norcliffe, Asifa Majid

Language, Volume 100, Number 1, March 2024, pp. 81-123 (Article)

Published by Linguistic Society of America

DOI: <https://doi.org/10.1353/lan.2024.a922000>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/922000>

VERBS OF PERCEPTION: A QUANTITATIVE TYPOLOGICAL STUDY

ELISABETH NORCLIFFE

University of Oxford

ASIFA MAJID

University of Oxford

Previous studies have proposed that the lexicalization of perception verbs is constrained by a biologically grounded hierarchy of the senses. Other research traditions emphasize conceptual and communicative factors instead. Drawing on a balanced sample of perception verb lexicons in 100 languages, we found that vision tends to be lexicalized with a dedicated verb, but that nonvisual modalities do not conform to the predictions of the sense-modality hierarchy. We also found strong asymmetries in which sensory meanings colexify. Rather than a universal hierarchy of the senses, we suggest that two domain-general constraints—conceptual similarity and communicative need—interact to shape lexicalization patterns.*

Keywords: perception verbs, colexification, lexical-semantic typology, sensory language, communicative efficiency

1. INTRODUCTION. In a seminal typological study, Viberg (1984, further developed in 2001) proposed that perception verb lexicons are universally constrained by a SENSE-MODALITY HIERARCHY, with vision at the top, followed by hearing, and then the lower senses of touch, taste, and smell. Viberg argued that the hierarchy is reflected in various asymmetries in sensory language, including asymmetries in the direction of semantic extensions of perception verbs, as well as in their relative token frequency, morphological complexity, and diachronic stability. The differential linguistic treatment of the senses has been claimed to reflect the fact that the higher senses of vision and, to some extent, audition dominate in human perception and neuroanatomy (Viberg 1984, 2001; see also Evans & Wilkins 2000). In this way, the perception domain is seen to have a ‘direct line’ to human biology, akin to the far better-studied domain of basic color terms, where similar biologically grounded universalist claims have been developed and debated (Berlin & Kay 1991, Gibson et al. 2017, Regier, Kay, & Khetarpal 2007, Roberson, Davies, & Davidoff 2000, Zaslavsky et al. 2019a).

Among the linguistic asymmetries attributed to sensory dominance, Viberg (2001, 2015) proposed that the hierarchy constrains the mapping between sensory meanings and words, with senses at the top of the hierarchy, especially ‘see’ and then ‘hear’, more likely to be lexically differentiated compared to those at the bottom. For example, a language might have a dedicated verb for ‘see’ and one or more perception verbs coexpressing the rest of the senses. Or it might have a dedicated verb for ‘see’, another for ‘hearing’, and then another verb covering the rest. According to Viberg, it is rare for languages to invert this pattern (e.g. with dedicated verbs for ‘smell’ and ‘taste’ and another verb covering ‘see’, ‘hear’, and ‘feel’).

* This project has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 836921 [LexPex]. Particular thanks to Harald Hammarström for sharing his digital collection of grammars and dictionaries. Many thanks also to Sebastian Sauppe and Claudia Mazzuca for their input at various stages, to Rik van Gijn for help with the Yuracaré data, and to the audiences at the 12th Dubrovnik Conference on Cognitive Science 2021, the Cross-linguistic Perspectives on Processing and Learning 2021, the 16th International Conference of Cognitive Linguistics 2023, and the Linguistics Association of Great Britain annual meeting 2023. Finally, we are grateful to Bodo Winter and Åke Viberg (who made themselves known as referees for this paper) and an anonymous referee for their substantial and constructive feedback on earlier drafts, and to John Beavers and Kristen Syrett for their editorial input throughout.

While Viberg's study is suggestive of a universal constraint on form-meaning mappings, it is not known whether the patterning he observed is typologically robust. As Viberg himself was careful to acknowledge (he considered his study a work in progress), his language sample was not balanced. Certain areas, such as North and South America and Oceania, were 'highly underrepresented', while European languages were overrepresented. Possibly, Viberg's proposed lexicalization universal is an artifact of imbalanced sampling. Aikhenvald and Storch (2013:3), for example, argue that '[t]here is hardly any doubt that universal claims concerning the preferred status of "vision" are highly Eurocentric, and do not hold for the majority of non-Western societies'.

Building on Viberg's classic study, we seek to establish whether there are typological regularities in the verbal encoding of sensory meanings and to shed light on the factors that shape perception verb lexicons across languages. We address the empirical limitations of previous work by drawing from a genealogically and geographically stratified sample of basic perception verb lexicons in 100 spoken languages. In the interest of reproducibility in linguistic typology (e.g. Corbett 2005, Haspelmath & Siegmund 2006, Song 2007), we replicate our analyses with data from CLICS³, the third installment of the Database of Cross-Linguistic Colexifications (Rzymiski et al. 2020). Across these two data sets, we examine whether there is evidence that lexicalization patterns are constrained by the proposed sense-modality hierarchy.

We also break new ground in considering whether domain-general factors influence the structuring of the perception domain across languages. Whereas Viberg focused on shared biology as a constraint on perception verb lexicons, other research traditions emphasize general conceptual and communicative factors as important in shaping how meanings are distinguished or COLEXIFIED (i.e. coexpressed) in words. These have never been considered fully in the context of perception verb typology; their relevance for other semantic domains suggests they are also of consequence here.

It is generally accepted that meanings packaged together in words tend to be conceptually related in some way (Brugman 1988, Evans 2010, Geeraerts 1997, Györi 2002, Lakoff 1987, Traugott & Dasher 2002, *inter alia*). This could be because a cognitive bias toward simplicity makes lexicons easier to learn and use (Xu, Duong, et al. 2020). Within a narrow semantic domain such as perception, we can presume a high degree of conceptual relatedness. It is possible, however, that some pairs of sensory modalities are conceptualized as more similar than others, resulting in asymmetric patterns of cross-modal colexification across languages. As we discuss in more detail in §2, Viberg's own work is suggestive of this, though it has never been systematically investigated with a balanced language sample. Nor is it clear how conceptual resemblance between senses—as a driver of colexification—would, in principle, interact with Viberg's proposed hierarchical constraint on lexical differentiation. From a different disciplinary perspective, anthropological research suggests that the conceptualization of the senses is highly variable cross-culturally (Classen 1997, Howes 2006), a view that would not lead us to expect global preferences in how sensory meanings combine in words. These diverging theoretical perspectives call for empirical investigation using quantitative methods.

While colexification can be a result of conceptual relatedness, it can be inhibited for communicative reasons. A long tradition of scholarship holds that languages efficiently trade off competing pressures of informativeness and simplicity (e.g. Carr et al. 2020, Gibson et al. 2019, Haspelmath 2021, Zipf 1949). A pressure for informativeness favors one-to-one mappings between meanings and words, while a pressure for simplicity favors colexification of related meanings. Lexicons appear to be efficiently

organized to balance these pressures. Word meanings have been shown to be more differentiated where a concept is frequently referenced, but to be broader in low-need parts of the domain (e.g. Gibson et al. 2017, Kemp & Regier 2012, Regier, Carstensen, & Kemp 2016, Xu, Liu, & Regier 2020). Lexical differentiation is also favored when there is a high chance of confusability in context, with colexification more likely in contexts where the intended meaning can be reliably disambiguated (Brochhagen & Boleda 2022).

To tease apart these various, potentially interacting, sets of constraints, we examine whether there are typologically robust lexicalization patterns in the perception domain using a combination of Bayesian regression modeling and semantic maps. First, with regression modeling, we test whether some sensory modalities are more likely than others to be encoded in dedicated perception verbs. Then, with weighted semantic maps (Georgakopoulos & Polis 2018, Haspelmath 2003), we examine whether there are asymmetries crosslinguistically in how sensory meanings combine in words.

To anticipate the results, we find—across two independent data sets—no evidence that lexical differentiation patterns in lockstep with the proposed sense-modality hierarchy. This challenges the assumption that a fixed, biologically grounded, hierarchy of senses is directly reflected in the lexical expression of perceptual categories. The results do, however, reveal global biases in colexification patterns (contra a radical cultural-relativist view). Certain pairs of sense modalities—in particular, hearing-touch, hearing-smell, and touch-taste—regularly combine in words across languages and geographical areas.

Strikingly, we find that smell and taste are rarely coexpressed in perception verbs. This is highly unexpected on conceptual grounds, given that these two senses are generally assumed to be closely aligned perceptually and conceptually (Small & Prescott 2005, Viberg 2001) and are associated in other lexical categories (Winter 2019). We present arguments grounded in domain-general principles for why smell and taste are kept apart in perception verbs, namely because of their potential for confusability in context.

Vision, finally, shows a robust tendency to remain lexically differentiated from all other senses. We argue that the proximal cause of this universal bias is also communicative, in this case related to frequency of use: across languages, vision verbs have a higher token frequency than nonvisual perception verbs (San Roque et al. 2015). The fact that vision is spoken of most frequently across cultures, could, in turn, be grounded in vision's dominance in human perception. As a distal cause of vision's observed lexical differentiation, this reprises—in a modified form—Viberg's original idea that vision stands apart from the other senses in its lexical encoding.

Taken together, we propose that the lexical structure of the perception domain is shaped by the interaction of two domain-general principles: conceptual relatedness (which provides the potential for colexification) and communicative pressures (which may inhibit it). This conclusion aligns with a growing body of evidence suggesting that lexicons evolve to strike a balance between simplicity and informativeness (Brochhagen & Boleda 2022, Gibson et al. 2017, Karjus et al. 2021, Kemp, Xu, & Regier 2018).

In §2 we discuss some preliminary working assumptions and flesh out Viberg's original proposal, as well as the contrasting conceptual and communicative theories of colexification. We examine lexicalization patterns in our balanced language sample in §3, beginning with a description of the database and sampling methods and then turning

to the results. We then replicate the analyses using CLICS³ with a different sampling method (§4). Finally, in §5, we interpret the findings in the context of the Vibergian sensory hierarchy in juxtaposition with cognitive- and communicative-oriented approaches to lexical-semantic typology.

2. THEORETICAL BACKGROUND TO THE CURRENT STUDY.

2.1. THE FIVE-SENSES MODEL OF PERCEPTION AND THE PROBLEM OF CROSSLINGUISTIC COMPARISON. Since Aristotle, Western thought has customarily ascribed to humans five basic senses: sight, hearing, touch, taste, and smell. In reality, as Winter (2019:12–14) cautions, the picture is more complicated. On the one hand, scientists recognize additional physiological systems for sensing pain, temperature, balance, and our body's location and movements (proprioception). On the other hand, complex interactions between senses call into question whether they should even be considered as separate. We could distinguish anywhere between three and thirty-three senses depending on whether we use type of stimulus (light = vision, chemical = smell and taste, mechanical = touch and hearing) or type of receptor (e.g. rods and cones in the eyes, sound and balance sensors in the inner ear, etc.) as the basis for identifying a perceptual sense (O'Meara et al. 2019). From a cross-cultural perspective, the notion of five basic senses has also been critiqued as Western-centric, with anthropological research indicating that this folk model is not universally relevant across cultures (Howes 1991).

Given this lack of consensus, how should we approach the linguistic expression of the senses? In a typological study, we can rely on the crosslinguistic data itself to delimit the set of sensory categories. On the ETIC approach to semantic typology (Evans 2010), we define the semantic distinctions within a domain based on the aggregate set of semantic distinctions lexicalized in the language sample. Using this maximally differentiated set of meanings, we can then make generalizations about the types of mappings found between lexemes and meanings. With this approach, we find that the five-senses model does, in fact, emerge at the typological level: the approach requires at least five senses, because there are languages (such as English, but also, as our study shows, many more; §3.1) whose perception verb lexicons differentiate between these meanings. Previous typological studies, including Viberg 1984 and Evans & Wilkins 2000, also tacitly rely on the etic approach and accordingly utilize a five-senses model.

While the etic approach justifies a minimum of five sensory categories, our study (like others before it) does make one simplifying assumption: we do not consider additional sensory categories that may also be lexicalized separately in verbs, for example, the perception of pain, temperature, proprioception, or interoception (internal bodily sensations). Geurts (2002) observes, for example, that in Anlo-Ewe (Atlantic Congo), the term *seselelame* refers to a generalized physical feeling in or through the body, as well as to the specific sense of a tingling in the skin and also certain emotional states. Evans and Wilkins (2000) report that internal bodily sensation is lexically distinguished from external touch in a number of Australian languages. Despite the relevance of such distinctions, too few sources mention them to make large-scale typological comparison of these meaning categories feasible.

Finally, a terminological note. Where a language uses a single lexeme to cover two or more meanings, we describe this neutrally as the COLEXIFICATION of meanings (following François 2008). The notion of colexification does not imply any particular emic characterization of the lexeme—whether it is vague (monosemous) or instead composed of multiple related but psychologically distinct senses (polysemous). We use the term CROSSMODAL colexification to refer to the colexification of multiple sensory meanings

in perception verbs. We also use the terms **MULTISENSE VERB** to refer to perception verbs that encode more than one sensory meaning and **UNIMODAL VERB** for perception verbs that encode a single sensory meaning. Having laid out these preliminaries, let us turn to the specifics motivating the current study.

2.2. SENSORY DOMINANCE AS A CONSTRAINT ON PERCEPTION VERB LEXICALIZATION. In his foundational typological study of perception verbs, Viberg (1984) proposed that perception verb lexicons are universally constrained by a biologically motivated sense-modality hierarchy. The hierarchy orders the senses along a fixed set of dominance relations, as shown in Figure 1.

sight > hearing > touch > {smell, taste}

FIGURE 1. Viberg's simplified sense-modality hierarchy.

Viberg's main evidence for the hierarchy came from asymmetries in the direction of semantic extensions of perception verbs. Based on a sample of fifty-three languages, Viberg proposed that semantic extensions proceed unidirectionally along the sense-modality hierarchy, such that a verb of seeing can take on the additional meaning of hearing, for example, or a touch verb can acquire a smell meaning, while the reverse direction of extensions apparently does not occur. Subsequent work by Evans and Wilkins (2000) with Australian languages supported Viberg's unidirectional generalization, and it has been characterized as a universal of lexical semantic change (Evans & Wilkins 2000, Hill 1988, Riemer 2010).

In this article, we are not directly concerned with Viberg's unidirectionality generalization but rather with the synchronic typological outcome of such processes of meaning extension. That is, regardless of the direction of the development (from *p* to *q* or *q* to *p*), are there synchronic regularities in which sensory meanings are packaged together in perception verbs across languages?

Confusingly, Viberg's unidirectional proposal has sometimes been misinterpreted as a generalization about which sense modalities are **MOST LIKELY** to extend their meaning, in other words, most likely to colexify. Huumo (2010:57), for example, writes that 'according to Viberg (2001) ... verbs of visual perception easily spread into other sensory meanings'. In a summary of Viberg's findings, Traugott and Dasher (2002:71) note the apparent parallel between vision's colexification behavior within and outside of the domain of perception, stating that 'sight and hearing, in particular, are subject to extension not only to other sense modalities but also to **INTELLECT**' (emphasis in original). More recently, Jędrzejowski and Staniewski (2021:6) write that Viberg's sensory hierarchy 'presupposes that a verb of seeing can be used to describe perception experiences belonging to the other sense modalities'.

In fact, Viberg made the opposite claim. He observed that 'in the majority of languages, "see" and "hear" are lexicalized as separate verbs' (2001:1297). In Viberg 2015 he articulated a stricter implicational characterization of this tendency for lexical differentiation, stating that:

[t]here are languages in which only 'see' is lexicalized as a simple verb. Other languages have 'see' and in addition a verb that covers all the non-visual modalities ... Other languages have 'see' and 'hear' and one or more verbs that cover the rest of the sense modalities. (Viberg 2015:109)

The confusion in the literature has probably arisen in part from Viberg's reference to vision's 'dominance' in semantic extensions. Viberg writes, for example, that "see" is

dominant in patterns of polysemy and has a tendency to extend unidirectionally to other sense modalities than sight' (2001:1307). If we have interpreted Viberg correctly, vision 'dominates' not because it extends its meaning within the perceptual domain more frequently than other modalities do, but rather because extensions never proceed upward from a 'lower' modality to vision.

To recap, Viberg proposed that the sense-modality hierarchy is reflected in the direction of semantic extensions (from higher to lower on the hierarchy) and in patterns of lexical differentiation (with verbs at the top of the hierarchy more likely than those beneath them to be lexically differentiated). In this article, we are concerned with the empirical validity of the lexical differentiation claim.

Beyond lexicalization patterns, Viberg proposed that the sensory hierarchy is reflected in other aspects of sensory language, including the relative frequency of perception verbs (from more to less frequent, in accordance with the hierarchy), their degree of formal complexity (Viberg 1984, 2001), and, more speculatively, their diachronic stability (with vision verbs more likely to be cognate within language families; Viberg 1984). Viberg (2001) considered these phenomena jointly as manifestations of 'lexical typological markedness' (in the sense of Croft 2002, Greenberg 1966), with vision exhibiting the least linguistically marked behavior and the lower senses the most marked behavior.

The markedness patterns have, in turn, been said to be 'firmly grounded in human biology and general cognition' (Viberg 2008:127), reflecting the relative dominance of the five senses in human perception (Evans & Wilkins 2000, Viberg 2001). In this way, the sensory hierarchy has a double ontological status, serving both as a description of typological patterns and as a functional explanation of those same patterns. Possibly, the differential amounts of neural tissue dedicated to each sense render aspects of perception 'intrinsically more accessible to consciousness and thus to language' (Majid et al. 2018:11369). Such a possibility receives some support from a recent study which finds that both the sensory salience of English words (as determined by subjective sensory norms) and the level of functional brain activation for the primary senses follow the Vibergian sensory hierarchy (Reilly, Flurie, & Peelle 2020). However, given the lack of consensus in the literature regarding how to even define the senses (§2.1), it should come as no surprise that not all researchers subscribe to the notion of a sensory hierarchy at the neuropsychological level (see Winter 2019 for discussion).

2.3. CONCEPTUAL RELATEDNESS AS A DRIVER OF COLEXIFICATION. While Viberg focused on the sense-modality hierarchy as an explanation for lexical patterning in the perception domain, other research suggests that domain-general conceptual factors could be at play. Many scholars claim that meanings grouped together in a single word tend to be conceptually related in some way (Brugman 1988, Evans 2010, Geeraerts 1997, Györi 2002, Lakoff 1987, Traugott & Dasher 2002, *inter alia*). Conceptual relatedness, in this context, refers to any kind of conceptual relation that could exist between meanings, including metaphorical mappings (x is a kind of y) as well as metonymic relations (x is a part of y). The colexification of conceptually related meanings could arise from a cognitive pressure for simplicity (Geeraerts 1997, Györi 2002, Xu, Duong, et al. 2020). Semantic memory may favor the colexification of meanings that relate easily to one another, thereby facilitating word learning, lexical retrieval, and interpretation (Ramiro et al. 2018, Srinivasan & Rabagliati 2015, Xu, Duong, et al. 2020). Typologically, the frequency with which two meanings colexify across unrelated languages is sometimes said to index the degree of naturalness or strength of the conceptual association between them (Evans 2010, Gentner & Bowerman 2009, Xu, Duong, et al. 2020, Youn et al. 2016).

Notably, Viberg's own work suggests that conceptual relatedness between sensory meanings could influence crosslinguistic colexification patterns. Viberg found that perception verbs extended their meanings unidirectionally, but he also observed that extensions did not always encompass contiguous meanings on the hierarchy (Fig. 1). So, for example, a verb of hearing could extend directly to smell without covering touch. Viberg captured the crosslinguistic patterns of semantic associations in his sample with a refined model (Figure 2). Evans and Wilkins (2000) found a similar pattern when focusing on perception verbs in Australian languages.

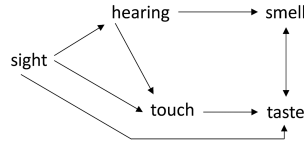


FIGURE 2. Viberg's refined sense-modality hierarchy for semantic extensions.

Figure 2 incorporates a number of generalizations. Sight does not associate directly with smell, hearing does not associate directly with taste, and touch does not associate directly with smell. Viberg (1984, 2001) speculated that some of the associations he uncovered reflect 'natural semantic relations' between sense modalities. He observed, for example, that the relationship between touch and taste on the one hand, and hearing and smell on the other, mirrors a distinction between contact and noncontact perception. In his sample, the inverse associations (hearing-taste and touch-smell) were not attested. He further proposed a close lexical relationship between smell and taste, reflecting perhaps that qualities of food sensed by olfactory receptors are experienced as flavors in the mouth. These observations suggest that conceptual relatedness may not apply to all pairs of senses equally.

There is a further wrinkle in this discussion of conceptual relatedness. To establish whether crossmodal colexification is driven by conceptual relatedness between sensory modalities, it is necessary to rule out the alternative possibility that multisense perception verbs develop via the diachronic process of SEMANTIC CHAINING. This is where a word semantically extends from meaning A to B, and then from B to C, and so on, along a chain of closely related meanings (e.g. Brugman 1988, Bybee, Perkins, & Pagliuca 1994, Hopper & Traugott 2003, Jurafsky 1996, Lakoff 1987, Ramiro et al. 2018, Xu, Duong, et al. 2020). Since sensory meanings belong to the same ontological domain and thus are closely associated, it may seem unlikely that semantic chaining is relevant. Recently, however, it has been proposed that crossmodal colexifications may come about via TRANSFIELD semantic extensions into the cognitive domain (Georgakopoulos et al. 2022, San Roque et al. 2018). So, a verb meaning 'see' extends its meaning outside of the perception domain to 'know' and then from 'know' back into the perception domain to 'hear'. In a study of colexification in the perception and cognition domains using data from CLICS², Georgakopoulos et al. (2022) found that associations between vision and the nonvisual modalities were almost always mediated by a cognitive meaning. To date, transfield semantic chaining has been implicated only for verbs of vision, possibly due to data sparsity for other sensory modalities. If it also applies to nonvisual verbs, it would undermine the idea that direct conceptual similarity between sensory modalities drives their colexification patterns. We consider this alternative possibility by examining patterns of both internal associations within sensory meanings and, separately, external associations with nonsensory meanings in verbs (see Evans &

Wilkins 2000, Georgakopoulos et al. 2022, Sweetser 1990, and Vanhove 2008a for detailed treatments of transfield semantic extensions of perception verbs).

At this point it is important to acknowledge that not everyone shares the presumption that there are universal conceptual associations between the senses. Sensory anthropologists have argued that perception, while a physical act, is mediated by cultural norms and values, and thus the conceptualization of the senses is highly variable cross-culturally (Classen 1997, Howes 2006). Ong (1967:3, cited in Howes 1991:26–27) writes, for example, that ‘cultures vary greatly in their exploitation of the various senses and in the way in which they relate their conceptual apparatus to the various senses’. Applied to crossmodal colexification, this view would not predict global regularities in mapping words to sensory meanings, although it may be compatible with the existence of areally bound patterns, on the assumption that colexification patterns are culturally transmissible (see Koptjevskaja-Tamm & Liljegen 2017 for an overview of areal tendencies in colexification).

Finally, Viberg did not discuss how crossmodal colexification driven by conceptual relatedness interacts with his proposed sensory-dominance constraint on lexical differentiation (§2.2). In at least one respect, they make the same prediction. On the sensory-dominance account, smell and taste would be the least likely of the senses to have dedicated verb forms (i.e. would be most likely to colexify) because they are the ‘lowest’ modalities on the proposed hierarchy. Smell and taste would also be predicted to share a verb form on the conceptual-relatedness account, given the close conceptual and perceptual association of these modalities. In other ways, however, the generalizations would appear to be in contradiction. As shown in Fig. 2, vision and hearing are both linked to other sense modalities, but according to the lexical differentiation claim, they should show a strong tendency to be lexically differentiated from other senses.

Part of the difficulty in interpreting these apparently contradictory generalizations may lie in the fact that Viberg did not report crosslinguistic frequencies. It could be the case that vision and hearing do link to other modalities, but only rarely. The apparent contradiction may also be the outcome of the fact that Viberg collapsed over two distinct kinds of lexical-semantic relations when formulating his graph of associations. In addition to polysemy proper (strict colexification),¹ Viberg included semantic extensions based on derivation/compounding (e.g. ‘hear’ + ‘odor’ > ‘smell’), in which a source perception verb (‘hear’) collocates with another element (‘odor’) to derive an ‘extended’ meaning (‘smell’). Following List (2023), we refer to this kind of process as PARTIAL COLEXIFICATION, because it involves partial identity of word forms. Because Viberg did not differentiate between strict and partial colexification, it is not clear from Fig. 2 whether some associations are instantiated by only one type of process. For example, Viberg’s two generalizations would be compatible if vision and hearing extend to the other senses via partial colexification, but remain lexically differentiated in strict colexification.

Altogether, these issues of interpretation highlight the fact that a definitive typology for the perception verb domain does not yet exist because potentially distinct processes have been conflated. In this article, we therefore focus exclusively on strict colexification, which allows us to establish the typological patterning that arises specifically from this lexicalization process. We examine whether the co-expression of sensory meanings in verbs can be explained by direct conceptual associations or through semantic chaining instead.

¹ It is perhaps worth clarifying that for strict colexification, Viberg did not rely on diachronic evidence to determine the directionality of extensions. Instead he argued that it was possible to identify a prototypical meaning for polysemous verbs based, for example, on how they were translated out of context. His contention was that the prototypical meaning was always higher on the proposed hierarchy than any secondary or extended meaning.

2.4. COMMUNICATIVE PRESSURES AS A DRIVER OF LEXICAL DIFFERENTIATION. Whereas conceptual relatedness encourages colexification, there may be communicative reasons for meanings not to colexify, even when a close conceptual association exists. Theories of communicative efficiency claim that language is shaped by a trade-off between competing pressures for simplicity and informativeness (Bentz et al. 2017, Carr et al. 2020, Fedzechkina, Jaeger, & Newport 2012, Gibson et al. 2019, Haspelmath 2021, Kemp & Regier 2012, Kirby et al. 2015, Levshina 2022, Piantadosi, Tily, & Gibson 2012, Smith 2020, Zipf 1949). At the lexical level, a maximally informative lexicon would have one distinct word per meaning. Such a system would, however, be difficult to learn. Learnability favors simpler lexicons that carve out larger regions of conceptual space—that is, it favors the colexification of related meanings.

One factor mediating between these pressures is communicative need, that is, how often speakers need to refer to a particular object or idea (e.g. Conway et al. 2020, Gibson et al. 2017, Karjus et al. 2021, Twomey et al. 2021, Zaslavsky et al. 2019b). It has been suggested that an efficient balance between simplicity and informativeness will lead to narrow categories (i.e. more lexically differentiated) in high-frequency portions of the semantic domain and to broader categories (i.e. colexified meanings) in low-frequency parts of the domain. This allows category systems to remain relatively simple while maximizing informativeness across the semantic space as a whole. Communicative need has been proposed to influence the lexical structuring of a variety of semantic domains, including color (Gibson et al. 2017, Zaslavsky et al. 2019a), kinship (Kemp & Regier 2012), numeral systems (Xu, Liu, & Regier 2020), and natural phenomena (Regier et al. 2016). It has also been implicated in the structuring of the English sensory lexicon. Winter, Perlman, and Majid (2018) found that across lexical categories, English has more words for visual concepts and that these are also more frequent than words for nonvisual sensory concepts. They argue accordingly that English perceptual vocabulary is oriented toward communicative need: vision, as the sensory category most frequently referred to, is also the one that is the most lexically elaborated.

If perception verb lexicons are geared toward communicative need across languages, then we would expect a typological tendency for vision to be lexically differentiated from the other senses. This is because vision verbs have a higher token frequency than nonvisual perception verbs not only in English, but also in all languages studied so far (Floyd, San Roque, & Majid 2018, Holmer 2021, Krishna, Arulmozi, & Mishra 2022, San Roque et al. 2015, Tchantourian & Vamling 2005, Viberg 1993, Winter et al. 2018), suggesting a common communicative need to talk about visual experiences across cultures. This dovetails with the first part of Viberg's sensory-dominance proposal (vision will be encoded with a dedicated verb because it is the dominant sense), but unlike Viberg's biological account, it identifies frequency of use as the proximal cause. Given that communicative need and Viberg's sensory-dominance proposal lead to the same prediction with respect to the lexical expression of visual perception, we return to the relative merits of these two approaches in the general discussion.

Another factor mediating between informativity and simplicity is context. Piantadosi et al. (2012) observe that where context is informative about meaning, unambiguous language is partly redundant and therefore inefficient. They argue that efficient communication systems will be ambiguous whenever context is informative about the thing being communicated. Conversely, two related meanings are less likely to colexify whenever context cannot reliably disambiguate. Supporting this idea, Brochhagen and Boleda (2022) found that colexification likelihood decreased across languages when

meanings were likely alternatives in context, and thus where confusability was particularly at stake. Brochhagen and Boleda observe that there may be a particular communicative disadvantage to expressing two meanings with the same word form when both meanings belong to the same ontological domain, because those meanings will appear in similar contexts and therefore cannot be reliably differentiated. This raises the possibility that colexification within a narrow semantic domain such as perception might be typologically disfavored overall or, intriguingly, that certain combinations of sensory meanings that are ‘too similar’ may tend to remain lexically differentiated from each other if they tend to be referenced in similar, nondifferentiating contexts. We return to the potentially disambiguating role of context later in the discussion (§5).

2.5. STRUCTURE AND SCOPE OF THE PRESENT STUDY. In this article, we ask if there are robust typological patterns to be found in this domain, and, to the extent that there are, we seek to isolate the factor(s) that best explain them. To do this, we organize our study around three interconnected empirical questions:

1. **UNIMODAL BIASES:** Are some sensory meanings more likely than others to be lexically differentiated?
2. **CROSSMODAL BIASES:** Do some combinations of sensory meanings tend to colexify in perception verbs more than others?
3. **TRANSFIELD SEMANTIC CHAINING:** Do crossmodal colexifications reflect first-order semantic associations (within the perception domain) or higher-order semantic associations (via semantic chaining with nonsensory meanings)?

From the pattern of results obtained, we assess which combination of factors—perceptual, conceptual, and communicative—most parsimoniously accounts for them.

We address the empirical limitations of previous work by drawing on a genetically and geographically balanced sample of basic perception verb lexicons in 100 spoken languages to create the PERCEPTION VERB DATABASE (PVDB). Manual data collection from published sources gave us more control over sampling than restricting ourselves exclusively to online databases, allowing balanced global coverage and enabling collection of full perception verb paradigms (spanning all five sensory meanings) for each language sampled. We also replicate our analyses with a separate language sample drawn from CLICS³. Unlike our own curated PVDB, CLICS³ is not balanced by family or area (e.g. no non-Pama-Nyungan languages of Australia are represented), and for perception verbs in particular, sensory meanings are not sampled equally across languages. It does, however, contain a sizeable set of nonoverlapping languages for which all basic sensory meanings are sampled (271 languages from forty-nine language families in total), allowing us to check the generalizability of our initial results. Notably absent in both of these samples are sign languages. Based on a convenience sample of twenty-four sign languages, Zeshan and Palfreyman (2019) conclude that perception verbs in sign languages do not show evidence of colexification because they follow the TRANSPARENCY PRINCIPLE OF SENSORY PERCEPTION, such that the place of articulation of a perception verb is linked to its meaning (e.g. ‘to see’ is articulated near the eyes, ‘to hear’ near the ears, etc.). This intriguing finding calls for systematic further study.

We limit our focus to the class of perceiver-oriented perception predicates (translational equivalents of *see*, *look*, *hear*, *listen*, etc., in which the perceiver role is realized as the grammatical subject in active clauses). We do not include stimulus-oriented perception verbs, in which the stimulus is linked to the subject role (e.g. *taste* as in *the soup tasted delicious*), as it became apparent during data collection that stimulus-oriented verbs are

often not lexicalized or documented for all senses (see Aikhenvald & Storch 2013, and Wälchli 2016 for a general critique of the copulative category of perception verbs; Viberg 2019 refines the definition and relates it to two other classes of phenomenon-based verbs).

We focus exclusively on instances of strict colexification as opposed to partial colexification (lexical-semantic associations based on derivation or composition; see §2.3). In this we diverge from Viberg (1984) and Evans and Wilkins (2000), who interchangeably used both as evidence of crossmodal associations. We believe that each phenomenon should be considered separately, as it is an empirical question whether strict and partial colexifications show the same behavior (see Münch & Dellert 2015 for some evidence that they do not). Procedurally, this means that if two sensory meanings are partially colexified (show partial formal identity), we treat them as separate perception verbs, because each verb form encodes a different meaning (see Norelliffe & Majid 2023 with regard to the partial colexification of perception).

3. LEXICALIZATION PATTERNS IN THE PERCEPTION VERB DATABASE.

3.1. THE DATABASE. The PVDB was constructed by hand for this study. It contains a genetically and areally balanced sample of basic perception verbs covering all five sensory meanings in 100 languages. In this section we describe the PVDB, outlining the sampling procedure, data sources, and criteria for data collection. We then give a brief overview of the types of perception verbs in the PVDB and present some initial typological generalizations, before turning to the main research questions. The database along with its metadata is available at https://osf.io/7kjp3/?view_only=2fdeaa9ac9bb4bbd9ac7a9c520e868e4.

SAMPLING. The languages represented in the PVDB were sampled following Miestamo, Bakker, and Arppe's (2016) method, in which each macroarea is sampled in proportion to its genealogical diversity (defined here at the family level). The goal is to minimize the possibility of areal bias, while maximizing genetic diversity. We stratified macroareas based on Hammarström and Donohue (2014), who partition the globe into six largely independent linguistic areas: Australia, Papunesia, North America, South America, Africa, and Eurasia. Each region has a history of linguistic contact within the area, but little historical contact with other areas. For our target sample size of 100 languages, we calculated the number of language families required for each area in order to preserve the relative proportions of families within areas, and we selected one language from each family (see the appendix for a list of languages).

We arrived at the sample size largely on the basis of availability and quality of documentation. Initial inspection of dictionaries revealed that a relatively comprehensive bilingual dictionary, grammar, or grammatical study that happened to refer to perceptual language was required in order to collect basic perception verbs spanning all five sensory modalities. For some areas, most especially Papunesia, many documented languages have only minimal coverage (e.g. a single word list or sketch grammar).² Lack of adequate sources prevented us from collecting data beyond about thirty Papuan languages (from thirty families), which put a cap on the total sample size.

Beyond constraining the overall sample size, the availability of language descriptions also complicated the task of sampling languages from each area and family (the so-called 'bibliographic bias'; see Miestamo et al. 2016 for discussion). Our method was to begin with a random sample of language families for each of the six macroareas.

² Swadesh lists, unfortunately, did not help us, as these include translational equivalents of only 'see' and 'hear' or, in the case of the extended Swadesh list, also 'smell'.

We then checked whether there was a language with a source of sufficient information about the basic perception verb lexicon. If not, we randomly selected an alternative family within that area. Where multiple languages within a family had adequate sources, we selected the language based on approximate geographical distance to other languages within that macroarea: we tried to sample equidistant languages (based on visual inspection of the map) within macroareas. This was not possible for all macroareas, due to the particular geographical spreads of language families. In Australia, for example, the bulk of the genealogical diversity (twenty-four of twenty-five language families) is found in a small corner of the northwestern part of the continent, while a single family, Pama-Nyungan, covers the remaining 90% of the continent.

We are aware that our sample likely retains other kinds of bias. As noted by Miestamo et al. (2016), ensuring that the genealogical diversity of each macroarea is equally represented may give rise to areal bias. For example, because Papunesia is genealogically very diverse, it is strongly represented in our sample (thirty languages out of 100). This makes it more likely that contact-based similarities between Papunesian languages may be overrepresented. By contrast, areal contact is probably less prevalent in our Eurasian subsample, given that Eurasia has comparatively less genealogical diversity and is represented in our sample by only eight languages. These eight languages are, moreover, spread over a greater land area than the Papunesian macroarea (Figure 3), decreasing likelihood of areal diffusion.

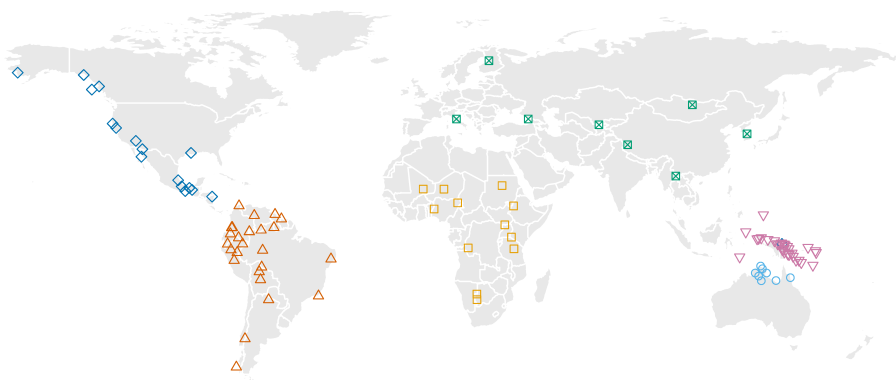


FIGURE 3. Geographic distribution of languages in the PVDB. Colored shapes indicate macroareas.³

It is often assumed that Australian languages are all related at a deep level (see Evans 2007), although, as Bowern (2010) observes, the comparative evidence for this remains scant. If this is the case, it makes Australia unique as the only continent with entirely related languages—though the time-depth of any Australian family (or ‘phylic family’, as Evans (2007:342) describes it), would be great, upward of 50,000 years. Many Australian languages were also spoken in contexts of extensive multilingualism, so borrowings are not uncommon (Evans 2007). As Bowern (2010) shows, untangling genetic relatedness from areal diffusion in this context is not trivial. For present purposes, we note formal resemblance between verbs of vision across Australian family-level groupings: Evans (1988:102) reconstructs the verb ‘to see’ for proto-Australian

³ This figure, like several others in the article, is presented in color in the electronic versions of this article, but in grayscale in the print version. Color versions of the figures are also available open access, along with the supplementary materials referenced throughout, at <http://muse.jhu.edu/resolve/233>.

as **na-* (**NHaa-* in proto-Pama-Nyungan), and *na-* (or complex forms based on *na-*) appears in six of eight languages in our sample. Restricting the Australian subsample to only those paradigms with no formal lexical resemblances would considerably reduce the genealogical representation of Australian languages. Given that our focus is on the mapping between verb forms and meanings and that our Australian subsample does not show uniformity at that level, we judged it preferable to retain the full subsample.

SOURCES. The PVDB data were collected manually from printed bilingual dictionaries and grammars (see the online supplementary materials S1 for a complete reference list). Reliance on published sources allows us to cast a wider typological net in comparison to consultation with native speakers, but inevitably has disadvantages. One issue is negative evidence: where a verb meaning was not listed in a source, we could not always determine whether this was due to a genuine lexical gap or a gap in documentation. We also did not have access to information about perception verb lexicons at the same level of detail across all languages (e.g. some languages had detailed information about the range of nonperceptual meanings associated with perception verbs, while others lacked this information).

CRITERIA FOR SELECTION OF MEANINGS. We collected lexical data for both noncontrolled experiencer meanings (translational equivalents of *see*, *hear*, etc.) and controlled activity meanings (translational equivalents of *look*, *listen*, etc.) for the five sensory modalities (sight, hearing, touch, taste, and smell).

The bilingual dictionaries we consulted were from the target language to various major languages: English, Spanish, French, Portuguese, German, and (in a couple of instances) Indonesian (for this last we had to back-translate entries into English). These languages functioned as the semantic metalanguages for the targeted verb senses. Some perception verbs in Romance languages are polysemous with respect to sensory meanings (the Spanish verb *sentir*, for example, can refer not only to tactile perception, but also to any other nonvisual modality and emotional experiences). Similarly, *feel* in English can refer to both tactile perception and emotion. Wherever possible, if a translation of a target term was ambiguous in the source with respect to the range of meanings encoded, we crosschecked against other sources to triangulate on possible denotations. We otherwise avoided sampling from languages for which the only available sources contained metalanguage ambiguities with respect to sense-modality encoding.

To identify basic perception verbs, we first looked up the target perception verb meanings in the source ('see', 'look', 'hear', 'listen', 'feel', 'touch', 'taste', and 'smell') and recorded the corresponding citation forms of the verbs in the target language, together with all additional translations of the form (perceptual and nonperceptual) that were listed. We then looked up each citation form in the target language and recorded all back-translations into the source language, in order to catch additional meanings that may not have been captured in the first entry. It was not uncommon for multiple perception verbs to be listed for a single modality meaning, without any obvious difference in meaning that would allow us to weight the selection of one over the other. In such cases we recorded all forms. An exception to this was cases where verbs showed grammatically encoded paradigmatic contrasts for particular categories, for example, object number (singular vs. plural) or object animacy (animate vs. inanimate). In such cases we selected only one member of the paradigm (the formally simplest one) as the representative verb form.

PERCEPTION VERB TYPES AND INITIAL TYPOLOGICAL GENERALIZATIONS. The PVDB contains a total of 664 perception predicates. Of these, 552 are morphologically simple lexemes, and 112 are morphologically derived or take the form of a compound expression.

A small number of languages lack a form for one or more sensory meanings: seven lack a form for ‘taste’, and one lacks a form for the tactile modality (‘touch’ or ‘feel’). For the visual and auditory modalities, where we can determine the event type of the predicate according to the translation in the source (‘see’ vs. ‘look’ and ‘hear’ vs. ‘listen’), three languages lack a verb form expressing the meaning ‘look’ (i.e. neither a dedicated verb for ‘look’ nor a verb that colexified ‘look’ and ‘see’), and five languages lack a verb for ‘listen’. For the tactile modality, it is not possible to determine the event type of the predicate with the same certainty, given that the English translation *feel* can refer to either the controlled activity of feeling or the noncontrolled experience, and sources frequently do not differentiate between these meanings (this is similarly the case for *taste* and *smell*). Bearing this ambiguity in mind, we can nevertheless state that for five languages, the only tactile perception verbs listed have an explicitly controlled activity meaning (i.e. these languages are missing a tactile verb with a translation compatible with a noncontrolled experience meaning). These absences are as likely due to gaps in sources as to genuine lexical gaps.

For the visual and auditory modalities, where we have information about the event type of the predicates, we checked the proportion of languages that lexically distinguish between agentive and experience event types vs. those that use a single verb to express both event types (i.e. for vision, the verb is translated as both ‘see’ and ‘look’). For vision, of the ninety-seven languages for which both event types are recorded in the source (recall from above that three languages are missing any kind of verbal encoding of ‘look’), thirty-five languages (36%) have separate verbs for ‘see’ and ‘look’, while the remaining sixty-two (64%) do not lexically differentiate between event types. We see a similar distribution for hearing verbs. Of the ninety-five languages for which both event types are recorded in the source, thirty-two (34%) have separate verbs for ‘hear’ and ‘listen’, while the remaining sixty-three (66%) do not lexically differentiate. For both vision and hearing, it is therefore more typologically common to collapse over event types than to lexically differentiate them, though neither strategy is rare.

Viberg (1984:137) observed for his language sample that perception verbs denoting controlled activities tended not to colexify multiple sense modalities. In line with this observation, multisense verbs that involve strictly controlled activity meanings are rare in the PVDB, with only three clear instances identified (with the caveat that we have reliable information about event-type distinctions only for sight and hearing verbs). In East Taa (Tuu), the verb *lk’ûle* is listed with the controlled activity meaning ‘to listen, to attend’ and also has the presumably experience meaning ‘to catch scent of’ (Traill 2018:125). (East Taa has a separate noncontrolled auditory perception verb *tâā*, which in addition to meaning ‘hear’ also encompasses the meanings ‘feel, taste, smell, perceive, understand’; Traill 2018:178). In Ma’di (Central Sudanic), the verb *mā*, which has the controlled activity meaning ‘to feel with hand’, also means ‘to taste, to try’ (Blackings 2000:67). In Leh Ladakhi (Sino-Tibetan), the verb *nyugches* means both ‘to search by touch, feel around with the hand’ and ‘to taste’ (Hamid 1998:95). Thus, while the typological usefulness of the controlled activity vs. noncontrolled experience distinction has sometimes been questioned since it is not made lexically in all languages (see e.g. Aikhenvald & Storch 2013), it appears to have a bearing on colexification potential.

A simple count reveals that multisense verbs are widespread crosslinguistically: multisense verbs appear in fifty-four of 100 languages in the database and are attested in all six macroareas (Figure 4). Of the remaining forty-six languages, forty-four use dedicated perception verbs to denote each of the five sense modalities; two languages are missing a ‘taste’ verb. It is also typologically common, then, for languages to lexically differentiate between all five sense modalities.

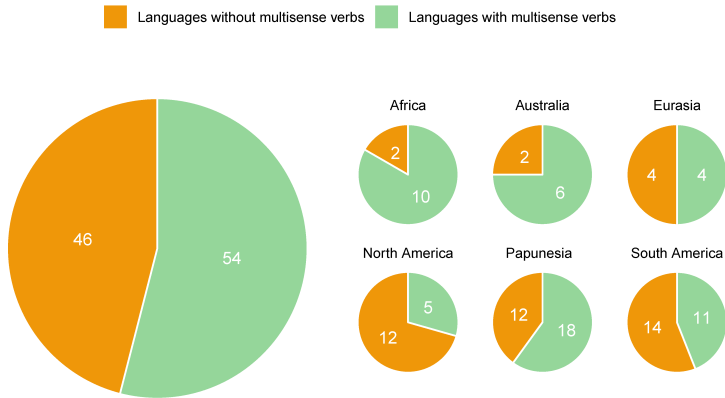


FIGURE 4. Proportion of PVDB languages as a whole and by macroarea with exclusively unimodal perception verbs vs. with one or more multisense verbs. Numbers indicate language counts.

Some example lexicons from the PVDB are provided in Figure 5 (these exclude trans-field meanings). Northern Uzbek and Tohono O’odham exemplify lexicons where there is a one-to-one mapping between sensory meaning and perception verb. The remainder show various types of crossmodal colexification (multisense verbs are highlighted in gray; some meanings are associated with multiple verb forms, which are separated by a comma).

Northern Uzbek (Turkic)	Tohono O’odham (Uto-Aztecan)	Sandawe (Isolate)
see ko’rmoq	see neith	see ǎǎ
look qaramoq	look ab neith	look ’ee
hear eshitmoq	hear kah	hear khé?e
listen tinglamoq	listen kaiham	listen khé?e
feel sezmoq	feel tahtk, tahtam	feel ǎǎ
touch tegmoq	touch tahtam	touch dzáá
taste totmoq	taste jehk	taste thomǐgé
smell iskamoq	smell hewagith	smell k’oótho

Central Kanuri (Saharan)	Tai Nüa (Tai-Kadai)	Mali (Baining)
see rúkin	see han ¹	see tlu
look rúkin, wúngin	look toi ² , meü ²	look binggi
hear fangín	hear ñin ²	hear tnari
listen fangín	listen faŋ ² , thoŋ ³	listen tnarimanas
feel fangín	feel ñin ² , ñom ² , tom ²	feel tnari
touch léngin	touch ñom ² , tom ²	touch tēk vēt
taste támbukin	taste tsim ²	taste tnari, tēs nēthēm
smell fangín, yínngin	smell lom ⁶	smell pērēn sēvēt

FIGURE 5. Examples of perception verb lexicons from the PVDB.⁴

⁴ References for the language examples in Fig. 5 are: Northern Uzbek: Butaev & Irisqulov 2008; Tohono O’odham: Saxton et al. 1983; Sandawe: ten Raa 2012; Central Kanuri: Cyffer & Hutchinson 1990; Tai Nüa: Luo 1999; Mali: Stebbins & Tayil 2012.

3.2. UNIMODAL BIASES: ARE SOME SENSE MODALITIES MORE LIKELY THAN OTHERS TO BE LEXICALLY DIFFERENTIATED?

DATA CODING. We created a variable **SENSORY MODALITY** (with five levels, Sight, Hearing, Touch, Taste, and Smell), which collapsed over agentive and experiencer meanings, and for each language coded the **LEXICALIZATION TYPE** for each sensory meaning. If a sensory modality in a given language was expressed by one (or more) multisense verbs, Lexicalization type was coded as ‘multisense’; if it was exclusively expressed by one (or more) unimodal verbs (i.e. there was no multisense verb associated with that modality), it was coded as ‘unimodal’.⁵ For the few cases where a sensory meaning had no associated verb form in a language (see §3.1), it was coded as ‘NA’.

ANALYSIS. The analysis was conducted with R 4.2.1 (R Core Team 2022). We used the package collection ‘tidyverse’ (version 1.3.2; Wickham et al. 2019) for data processing and visualization and ‘brms’ (version 2.17.0; Bürkner 2017, 2018, 2021) for Bayesian mixed-effects regression models. Analysis code is available at: https://osf.io/7kjp3/?view_only=2fdeaa9ac9bb4bbd9ae7a9c520e868e4.

To determine whether some sensory meanings are more likely to be expressed by unimodal verbs compared to others, we used Bayesian logistic regression. In the model, the dependent binary variable was Lexicalization type (whether a language exclusively used a unimodal verb vs. multisense verbs), and the sole predictor was Sensory modality (Sight, Hearing, Touch, Taste, and Smell). This predictor was Helmert-coded. Helmert coding compares each level of a categorical predictor variable to the mean of subsequent levels of that variable. This allowed us to compare (i) Smell vs. Taste, (ii) Touch vs. Smell and Taste, (iii) Hearing vs. Touch, Smell, and Taste, and (iv) Sight vs. Hearing, Touch, Smell, and Taste. With Helmert coding, we can determine whether languages are more likely to exclusively use unimodal verbs to encode sensory meanings increasingly higher on the proposed modality hierarchy. It also allows us to check whether, as predicted by theories of communicative need, visual meanings are more likely to be lexically differentiated compared to nonvisual sensory meanings.

To allow macroareas to have different baseline preferences for multisense verbs, we included a random intercept for Macroarea. Since it is also possible that contact patterns may have differing effects on colexification in different areas, we included by-Macroarea random slopes for Sensory modality. Because the PVDB samples exactly one language per language family, random effects for language family were not required.

We placed weakly informative, normally distributed priors on population-level (‘fixed’) slopes ($\mu = 0$, $\sigma = 1$; Lemoine 2019, McElreath 2020). We used the default brms priors for the standard deviation of group-level (‘random’) effects and residual errors, which take only positive values (Vasishth et al. 2018:150). We report whether the 95% credible intervals (CrIs) of the posterior distributions for each predictor included zero and interpret those that do not contain zero as providing the strongest evidence for the effect of the predictor on the dependent variable. We also report the posterior probability of the effect being above or below zero.

⁵ According to this coding scheme, when a language is recorded as having both a unimodal and multisense verb to express a given modality (e.g. Central Kanuri forms for ‘smell’; Fig. 5) it is coded as ‘multisense’. If we instead code such cases as ‘unimodal’, the same pattern of results holds (see supplementary materials S2).

Markov chain Monte Carlo (MCMC) sampling was performed with 4,000 iterations each for four chains (2,000 iterations discarded after warmup). To ensure convergence, we set the drift parameter delta to 0.99. Visual inspection of the chains and Rhat values of 1.0 indicated that the model converged. Posterior predictive checks showed that the model fit the data well.

RESULTS. The log odds of a language encoding the visual modality exclusively with a unimodal verb were reliably higher than for nonvisual modalities (log odds: 0.50, $SE = 0.14$; 95% CrI [0.25, 0.79]), with a very low posterior probability of the effect being negative ($\beta_1 < 0 = 0.2$). The comparison between Hearing and Touch, Taste, and Smell revealed that languages were LESS likely to encode hearing meanings with a unimodal verb compared to the lower modalities (-0.12 , $SE = 0.08$; 95% CrI [-0.27, 0.05]), though the evidence for this effect was somewhat less robust, with a small number of posterior samples above zero ($\beta_1 > 0 = 6.2$). Languages encoded Touch meanings reliably less often with unimodal verbs compared to Smell and Taste (-0.34 , $SE = 0.13$; 95% CrI [-0.60, -0.10]; $\beta_1 > 0 = 0.8$). There was no reliable difference between Smell and Taste (0.08, $SE = 0.30$), with the posterior distributions firmly overlapping with zero (95% CrI [-0.48, 0.73]); see Figure 6.

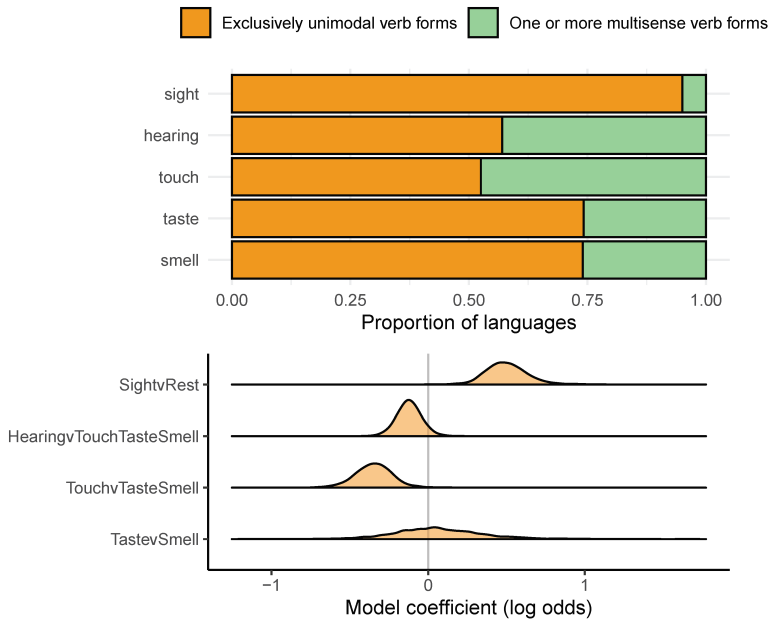


FIGURE 6. Top: proportion of languages that encode each sensory modality with unimodal verb(s) exclusively vs. one or more multisense verbs. Bottom: posterior distribution of logistic regression coefficients predicting whether languages have exclusively unimodal verbs vs. one or more multisense verbs.

As an additional simple check that unimodality is not hierarchically constrained within languages (as proposed by Viberg 2015), we counted the number of languages with basic perception verb lexicons consistent with the hierarchy and those that inverted

this ordering.⁶ Of the forty-seven languages containing no lexical gaps for sense modality, only nine (19%) conformed to the hierarchy (i.e. if a modality was expressed exclusively with a unimodal verb, then so were modalities above it on the hierarchy).

In sum, while there is a strong bias for vision verbs to be unimodal, we did not find descending rates of unimodality across the senses in lockstep with the proposed sense-modality hierarchy, when we either aggregate over languages or check for implicational tendencies within them.

3.3. CROSSMODAL BIASES: DO SOME COMBINATIONS OF SENSE MODALITIES COLEXIFY MORE THAN OTHERS? Having established that there are typologically robust asymmetries between visual and nonvisual senses, we ask whether some combinations of sensory meanings tend to be coexpressed in a word more than others. To investigate this, we make use of SEMANTIC MAPS, a method for visually representing typological patterns of linguistic coexpression (for overviews, see Croft 2022, Georgakopoulos & Polis 2018, and Haspelmath 2003).

A BRIEF OVERVIEW OF SEMANTIC MAPS. A semantic map is a graph in which nodes indicate meanings or functions (e.g. sensory meanings) and edges indicate associations between meanings (i.e. the coexpression of two meanings in a single perception verb lexeme). Viberg's original network of semantic extensions (Fig. 2) can be considered an early example of a semantic map. Because Fig. 2 represents the direction of semantic shifts (as inferred from either morphological derivation or identification of a prototypical vs. secondary meanings), it is a 'directed' graph (the edges have arrow heads conveying the direction of semantic extension). In this study we use undirected graph representations because we do not have information about historically prior states of verb meanings for most languages in our sample. We therefore do not consider the directionality of semantic shifts, but simply synchronic patterns of associations between sensory meanings.

The construction of a semantic map proceeds according to the principle of CONNECTIVITY (Croft 2001:96, 2002:134): the graph should be constructed so that every lexeme is represented by a connected region of the graph. Importantly, the graph should obey MINIMALITY by including only the minimal number of edges required to guarantee connectivity for each lexeme (Croft 2022:72, Georgakopoulos & Polis 2018:6, Zwarts 2010:379).⁷ For example, in Figure 7, a node is required between A and B in order to capture form 1, and a node is required between B and C to capture form 3. This connected set of edges incidentally also captures form 2; an edge between A and C is vacuous and therefore is not included. Another way of expressing the idea of minimality is that an edge should be added between nodes only if there is a language that INDEPENDENTLY COLEXIFIES the meanings represented by those nodes, that is, without also colexifying intervening nodes (we borrow this term from Gast and Koptjevskaja-Tamm (2022)).

⁶ For this calculation we treated instances of complex perception predicates derived from other perception verbs (e.g. verb+noun collocations such as hear+smell_[N], meaning 'smell') as 'unimodal' because the complex predicate encodes only a single modality meaning, so this differs from multisense verbs proper. It is unclear whether this is how Viberg intended his claim to be interpreted, as he did not fully differentiate between patterns arising from strictly colexified multisense verbs and those arising from partially colexified forms (see §2.3).

⁷ Semantic-map graph structures differ from less-constrained graph structures sometimes used by typologists in which every pairwise colexification relation is represented by an edge (see the web-interface version of CLICS and papers based on CLICS data (e.g. Jackson et al. 2019)). Croft (2022) refers to such graphs as 'pairwise co-expression graphs'.

	Meaning A	Meaning B	Meaning C
Form 1	1	1	0
Form 2	1	1	1
Form 3	0	1	1

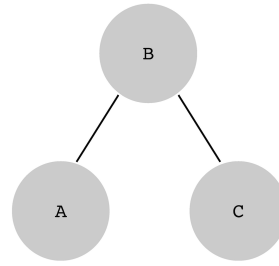


FIGURE 7. Example table of meanings and forms and its corresponding semantic map.

As a result of the connectivity principle, semantic maps can be interpreted as a set of implicational statements about the possible combinations of meanings associated with a form (Haspelmath 1997): if two nodes are not directly connected, this implies that forms will express both of those meanings only if they also include the meanings represented by the intermediate nodes. The immediacy of the connection between meanings in turn is taken to be informative about the relative degree of similarity between those meanings (Zwarts 2010:379). In Fig. 7, A is closer to B than it is to C, and therefore A and B are taken to be more semantically similar than A and C. More generally, the distance between two meanings in a graph can be defined as the length of the shortest path connecting them (Zwarts 2010:379).

In addition to summarizing synchronic semantic relations between lexicalized meanings, semantic maps allow us to make inferences about historical processes of semantic change. Because extension processes are motivated by similarity between meanings, it follows that semantic extensions will proceed stepwise along the network of meanings on a semantic map (Haspelmath 1997). So, although undirected maps cannot tell us about the directionality of an extension (A to B vs. B to A), they do allow us to discern whether an extension between A and B is direct or is the outcome of semantic chaining (§2.3). In the context of the present study, we can further distinguish between INTRAFIELD SEMANTIC CHAINING (involving stepwise extensions within the perception domain) and TRANSFIELD SEMANTIC CHAINING (involving extensions into and out of nonperception domains; see §3.4 and §4.4).

CONSTRUCTING WEIGHTED SEMANTIC MAPS OF CROSSMODAL COLEXIFICATIONS. To construct a crossmodal semantic map, we extracted the set of multisense perception verbs from the PVDB, which consisted of sixty-two multisense verbs from fifty-four languages and fifty-four language families. Fifteen distinct types of multisense verbs are represented in total, differing from each other with respect to the specific number (range 2–4) and type of sensory meanings encoded (see supplementary materials S3).

We transformed the multisense data into a matrix, with sensory meanings listed by column and verb forms by row (see Fig. 7). Each cell received a 1 if the verb form encoded the associated sensory meaning, and a 0 otherwise. We then computed the minimal number of edges required to generate the underlying graph structure, using a Python script developed by Regier, Khetarpal, and Majid (2013:93–94). We used the R package ‘ggraph’ (version 2.0.5; Pedersen 2021) to plot the set of edges as an undirected graph using the Fruchterman-Reingold layout algorithm.⁸

⁸ This algorithm belongs to a set of commonly used force-directed algorithms, which model a graph as a physical system where nodes are attracted and repelled according to some force. The aim of these algorithms is to lay out the graph in a way that optimizes readability and allows the user to easily perceive its topological

Following recent approaches (Georgakopoulos et al. 2022, List, Terhalle, & Urban 2013), we apply several analysis metrics to increase informativity. First, crosslinguistic frequencies of pairwise crossmodal colexifications are represented as edge weights: the thicker the edge, the greater the number of connections between nodes (i.e. the more often the two meanings are expressed together in a single verb across languages). We also use a node metric—CENTRALITY—to measure how important a given node is in the network. A node's DEGREE CENTRALITY is simply the sum of the edges connected to that node. A high score means a node has a larger-than-average number of connections. In weighted networks, the STRENGTH score for a node (also referred to as WEIGHTED DEGREE) sums the weights (i.e. crosslinguistic frequency) of all edges connected to that node. For example, a node with two edges, each with a weight of 2, has a strength score of 4 ($2 + 2$); a node with four edges, each with a weight of 1, would also have a strength score of 4 ($1 + 1 + 1 + 1$).

Edge weights indicate the observed crosslinguistic frequencies of crossmodal colexifications, but do not say anything about the overall statistical likelihood of the coexpression of different pairs of sensory meanings. We therefore additionally compute the coexpression probabilities between sensory meanings using a simple analytic model based on the probability mass function of the hypergeometric distribution and implemented in the R package 'co-occur' (Griffith, Veech, & Marsh 2016). The model, which is mathematically equivalent to Fisher's exact test (Arita 2016), is borrowed from ecology research. Using Griffith et al.'s cooccurrence model, we can determine rates of colexification that differ significantly from expected values by calculating the probability of finding a number of coexpressed meanings equal to or more extreme than the observed counts. These probabilities can be directly reported as *p*-values, without reference to a statistic. The calculations are implemented using the function `cooccur()`, which takes as input a matrix of meanings by verb forms and returns, for each meaning pair, the probabilities that those meanings could be colexified less than or greater than what was observed in the data, thus providing significance levels for negative and positive association patterns at a specified alpha level (we use $\alpha = 0.05$). Using the output of these models, we color edges on the maps according to whether colexifications are significantly negatively or positively associated or are neutral (i.e. observed neither reliably more nor less than expected values). See supplementary materials S4 for further details about Griffith et al.'s (2016) cooccurrence model.

A WEIGHTED SEMANTIC MAP OF CROSSMODAL COLEXIFICATIONS. Figure 8 presents the semantic map of crossmodal colexifications observed in the PVDB. The map reveals that hearing and touch are the most central nodes (with strength scores of 75 and 80, respectively). Taste and smell have lower strength scores (45 and 49, respectively), and sight has the lowest (7). That is, sight is the least central node in the network. Reflecting its unimodal bias, vision shows only a weak relationship with the other modalities. It is directly linked to hearing and touch, but not to taste or smell. In diachronic terms, this implies that sight will not semantically extend to smell or taste (or smell/taste will not extend to sight) without first extending to hearing or touch. Furthermore, the orange edges indicate that sight is negatively associated with touch and hearing, meaning that these associations appear significantly less often than their expected value (see model output, supplementary materials S5).

structure (e.g. by reducing the number of edge crossings and maintaining a uniform edge length). For the intrafield semantic maps presented below (e.g. Fig. 8), the choice of layout algorithm does not make a difference to the readability of the graph, given the small number of nodes overall. For the more complex transfield maps (Figs. 9 and 12), the choice of algorithm is more important.

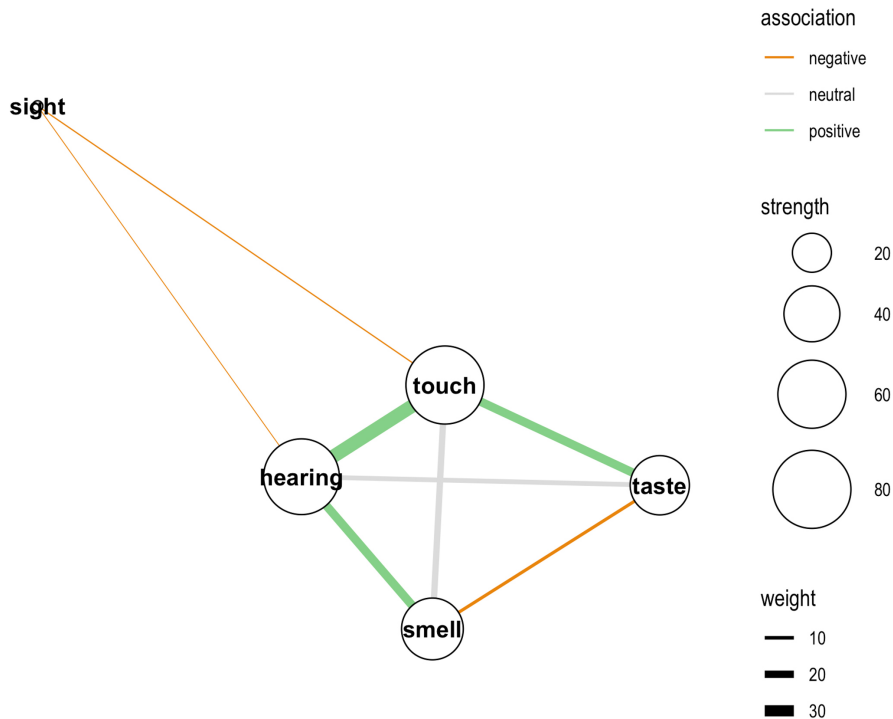


FIGURE 8. Weighted semantic map of crossmodal colexifications in the PVDB. Nodes indicate meanings, with size reflecting strength scores. Edges indicate colexifications, thickness reflects crosslinguistic frequency of colexified meanings, and colors indicate whether meanings are significantly negatively (orange), positively (green), or neutrally (gray) associated, according to the pairwise probabilistic coexpression analysis.

Given the overall typological rarity of crossmodal colexifications involving visual meanings, we return to the original data for further qualitative exploration. Only four languages in the PVDB have a multisense verb that coexpresses a visual meaning (see Table 1), but these languages are attested in four different macroareas, so despite their overall rarity, this is not a geographically localized phenomenon.

LANGUAGE	FAMILY (AREA)	VERB FORM	SENSORY MEANINGS
Kuku-Yalanji	Pama Nyungan (Australia)	nyajil	‘see, hear’
Sandawe	Isolate (Africa)	lāā	‘see, feel’
Mayangna	Misumalpan (North America)	dakanin	‘see, hear, listen, feel’
Kalam	Nuclear Trans New Guinea (Papunesia)	nŋ	‘see, hear, listen, feel, touch, smell’

TABLE 1. Crossmodal colexifications that include visual meanings in the PVDB.

The edge between sight and hearing is supported by a single language, Kuku-Yalanji (a Pama-Nyungan language spoken in the region around Cape York, Northern Queensland), whose multisense verb *nyajil* independently colexifies ‘see’ and ‘hear’ (Hershberger & Hershberger 1982:135). Evans and Wilkins (2000) note that sight > hearing meaning extensions (including extensions based on derivation) may be an areal phenomenon in this region. The edge between sight and touch is supported by

Sandawe (language isolate), whose multisense verb *lãã* independently colexifies ‘see’ and ‘feel’ (ten Raa 2012:168). The Mayanga (Misumalpan) verb *dakanin* colexifies ‘see’ together with both ‘hear/listen’ and ‘feel’ (von Houwald 1980:86). Finally, Kalam (Trans New Guinea) has a multisense verb *ny*, which coexpresses all sense-modality meanings except for ‘taste’. We can be confident that the absence of a taste meaning is not due to lack of documentation in this case—Pawley (2020) has written extensively about the semantics of this verb. In addition to denoting basic perception meanings, it also expresses a broad range of perceptual and cognitive meanings. The Kalam verb *ny* is the only perception verb attested in the PVDB that coexpresses smell and sight, emphasizing the typological rarity of this combination of meanings. The combination of sight and taste is even rarer still: no language in the PVDB coexpresses these meanings (either directly or via intrafield chaining).

Inspecting the nonvisual senses in Fig. 8, we see that they are all directly linked with each other. The statistical likelihood of colexification differs among pairs of meanings, however. Three pairs—hearing-touch, hearing-smell, and touch-taste—are significantly positively associated. For hearing-taste and touch-smell, colexification is reliably neither under- nor overrepresented. Smell and taste, finally, are significantly negatively associated. This pair of modalities is coexpressed in multisense verbs in nine languages. Critically, in only one of these, Kuku-Yalanji, are smell and taste independently colexified in a verb (i.e. to the exclusion of other sensory meanings). The fact that smell and taste tend to be colexified in the company of other sensory modalities implies that intrafield semantic chaining drives their coexpression.

To summarize, visual meanings rarely colexify with the nonvisual senses and never with taste. Of the nonvisual pairs, hearing-touch, hearing-smell, and touch-taste colexify more frequently compared to hearing-taste, touch-smell, or taste-smell, though direct links between all sensory meanings are observed in the data. Taste-smell colexifications are significantly underrepresented and seem to arise primarily via semantic chaining with other sensory meanings.

3.4. TRANSFIELD SEMANTIC CHAINING: ARE CROSSMODAL COLEXIFICATIONS MEDIATED BY NONSENSORY MEANINGS? In this section we take the semasiological step of considering all meanings expressed by perception verb forms in the PVDB, in order to determine whether crossmodal colexifications tend to be the outcome of regular processes of semantic chaining involving nonsensory meanings (Georgakopoulos et al. 2022, San Roque et al. 2018). For convenience, we refer to meanings not included in the set of basic sensory meanings as TRANSFIELD meanings. For purposes of comparison, we standardized the transfield meanings across languages and treated closely related synonyms as a single meaning (see supplementary materials S6). We then take the set of intrafield and transfield meanings associated with perception verbs in the PVDB and again infer a semantic map, following the same procedure described in §3.3.

In addition to calculating the weighted centrality of nodes, following the approach of Georgakopoulos et al. (2022), we also applied a community-detection algorithm to determine the community structure of the graph.⁹ A community is a collection of nodes with relatively few connections to other groups of nodes, but a high number of connections among themselves. In the present study, where nodes represent verb meanings, a community corresponds to a cluster of meanings closely connected to each other.

⁹ We did not use community detection for the intrafield semantic map (Fig. 8) due to the small number of nodes overall.

Identifying such substructures in the network can establish whether different sensory modalities are associated with distinct clusters of meanings. For community detection, we used the spinglass algorithm, which has been shown to yield accurate results for small networks of fewer than 1,000 nodes (Yang, Algesheimer, & Tessone 2016). For the layout, we used the Davidson-Harel algorithm as implemented in the R package ‘igraph’ (version 1.3.3; Csárdi & Nepusz 2006), which generated a readable topography with evenly distributed nodes and few edge crossings.

To preserve the readability of this larger map and to reduce the chance of including spurious homonyms, we follow the standard practice of pruning the edges of the maps (Croft 2022). We include edges only between meanings with a frequency of two or more. The final map, given in Figure 9, is based on ninety-nine languages and 340 verb forms.

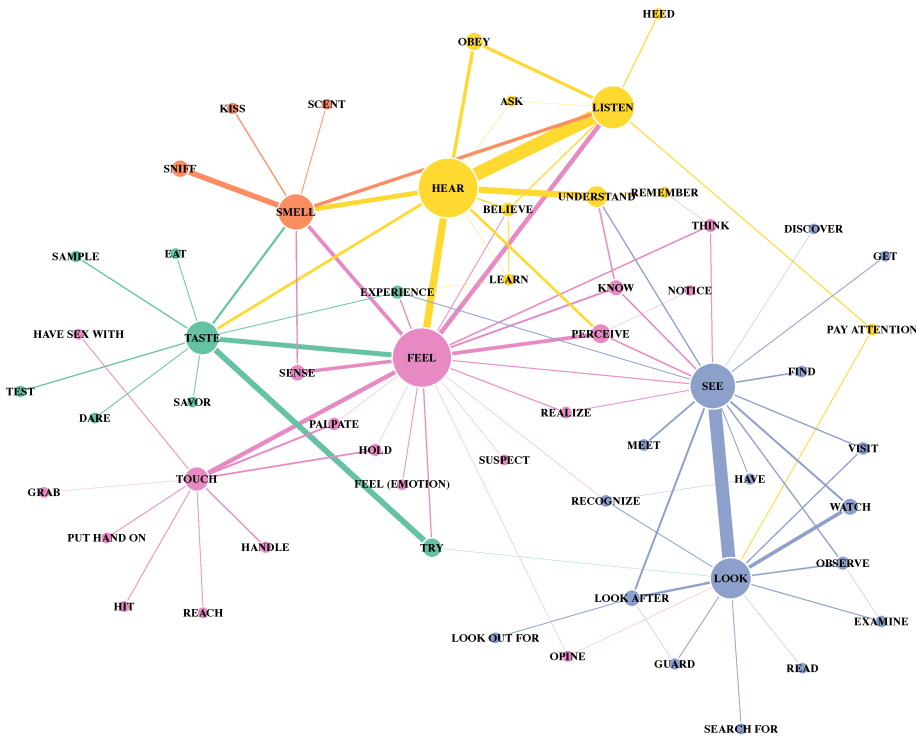


FIGURE 9. Weighted semantic map of meaning associations based on colexification data in the PVDB (seed concepts: ‘see’, ‘look’, ‘hear’, ‘listen’, ‘feel’, ‘touch’, ‘smell’, ‘taste’). Thickness of edges indicates relative frequency of colexified meanings, size of nodes indicates strength scores, and colors represent different communities detected within the structure.

Five communities were detected, each clustered around a sense modality, indicating distinct sets of transfield meanings. There are, however, transfield meanings that bridge sense modalities. Most significantly, the visual and auditory modalities are connected only via intermediary nodes: ‘see’ and ‘hear’ are linked via the general perception term ‘perceive’ and the cognitive meaning ‘understand’, while ‘look’ and ‘listen’ are linked by the focused attention meaning ‘pay attention to’. The lack of a direct edge between visual and auditory meanings is consistent with previous proposals that cross-modal colexifications involving vision arise via initial extension to a transfield meaning

(Georgakopoulos et al. 2022, San Roque et al. 2018). Figure 9 suggests that there are, in fact, multiple intermediary nodes between sight and hearing, suggesting distinct pathways of semantic extension.

There is a low-weighted edge between ‘see’ and ‘feel’, but these meanings are also linked by multiple nonsensory meanings, for example, ‘realize’, ‘know’, and ‘think’, which suggests that see-feel colexifications are typically the result of transfield semantic chaining. It is not possible to determine whether the direct link between ‘see’ and ‘feel’ represents a data gap (i.e. is due to lack of representation of transfield meanings in the sources for the verbs in question), but this certainly raises that possibility.

While vision shows only a tenuous connection to the other senses, there is stronger support for direct links between the nonvisual modalities, suggesting that nonvisual cross-modal colexifications arise from extensions within the perception domain. This finding is consistent with Georgakopoulos et al.’s (2022) earlier observation based on CLICS² data. Some directly connected pairs of senses are also connected via intermediate transfield nodes (e.g. ‘experience’ connects ‘taste’ to ‘feel’), which may point to additional pathways of semantic extension. This is not the case for all modalities, however, most especially for smell, which is associated with few transfield meanings overall; those that are attested more than once are restricted to meanings very closely associated with olfaction or perception: ‘to sense’, ‘to sniff’, and ‘to scent’, as well as ‘to kiss’ (see Schapper 2019 for a comprehensive study of smell-kiss colexifications in Southeast Asia). In sum, vision is only tenuously directly linked to the nonvisual senses (showing a direct link only to ‘feel’), while the nonvisual senses show evidence of direct links to each other.

3.5. SUMMARY. Overall, then, the analyses of the PVDB shed new light on prior claims about the coding of sensory meanings in verbs. Consistent with the first part of Viberg’s sensory-dominance proposal as well as with theories of communicative need (but contrary to generalizations in the secondary literature; e.g. Huumo 2010, Jędrzejowski & Staniewski 2021, Traugott & Dasher 2002), a logistic regression model revealed a strong bias for vision to be lexicalized with a dedicated verb. In contrast to the sensory-dominance proposal as a whole, the unimodal bias did not extend to audition, and there was no tendency for lexicalization rates to be in lockstep with the proposed sensory hierarchy.

Not all combinations of sensory meanings were equally likely to colexify. A semantic map enriched with pairwise coexpression probabilities revealed that all nonvisual modalities were directly linked to each other, but only three pairings—hearing-touch, hearing-smell, and touch-taste—were significantly positively associated. The higher frequency of hearing-smell and touch-taste relative to hearing-taste and touch-smell is consistent with Viberg’s proposal of a bifurcation between the senses based on contact. Such a generalization does not explain the high frequency of hearing-touch colexifications, however; to the best of our knowledge this strong typological tendency has not been discussed elsewhere in the literature, and we consider the conceptual basis for this association in the discussion (§5). Viberg also proposed that smell and taste were closely associated. Strikingly, we found no support for this proposal: smell and taste were significantly negatively associated, and there was only a single attested instance of direct colexification.

Finally, a semantic map of transfield colexifications revealed that almost all associations between sense-modality meanings were direct (first-order) and not the result of semantic chaining via transfield semantic extensions. Vision alone showed exceptional behavior in being linked to audition exclusively via nonsensory meanings. Before turning to a deeper discussion of these findings, we consider the generalizability of these results with another database, CLICS³.

4. LEXICALIZATION PATTERNS: A REPLICATION WITH CLICS³. As questions about reproducibility in linguistic typology are increasingly raised (e.g. Corbett 2005, Haspelmath & Sigmund 2006, Song 2007), we replicated our analyses with an existing database, CLICS³, the latest iteration of the Database of Cross-Linguistic Colexifications (Rzymiski et al. 2020). We downloaded the SQL database from the CLICS³ GitHub repository and preprocessed the data as follows: lexical forms with glosses and meta-data were extracted, converted to CSV format, and imported into R. We then filtered to include only verb forms with a basic perceptual meaning (any of the CLICS³ concepts ‘see’, ‘look’, ‘hear’, ‘listen’, ‘smell (perceive)’, ‘smell’, ‘feel (tactually)’, ‘feel’, ‘touch’, or ‘taste (something)’).

4.1. UNIMODAL BIASES IN CLICS³.

SAMPLING, CODING, AND ANALYSIS. CLICS³ contains a large number of languages (3,156 in total), but it has very uneven coverage of the five sense-modality meanings. While 2,541 language varieties have a lexical entry for the concept ‘see’ and 2,244 for ‘hear’, only 1,369 languages have one for ‘smell (perceive)’, 429 for ‘feel’ or ‘feel (tactually)’, and 620 for ‘taste (something)’. For the unimodality analysis, we can only draw from languages for which all five sense-modality meanings are sampled. This is because, in order to determine whether a verb recorded in CLICS³ is unimodal, we must rely on negative evidence: we code a verb form as ‘unimodal’ if it is exclusively associated with a single sensory concept, and we can do this only for languages for which all five sensory concepts are sampled (see Gast & Koptjevskaja-Tamm 2022 for a similar approach to extracting information about lexical differentiation from CLICS). We therefore filtered the CLICS³ perception verb data to include only languages with coverage of all sense-modality meanings. Doing this resulted in a much smaller data set of 287 languages from fifty-eight families. We then excluded a further sixteen languages that were also represented in the PVDB, yielding a final data set of 271 languages from forty-nine families. Note that, while the sample of languages is larger than the PVDB, it is not balanced by family or area.

We created a SENSORY MODALITY variable, and coded and analyzed the data as described in §3.2. The CLICS³ model followed the same specifications as the PVDB model, with the addition of a random intercept and slopes for LANGUAGE FAMILY, given the structure of the data. Visual inspection of the chains and Rhat values of 1.0 indicated that the model converged. Posterior predictive checks showed that the model fit the data well.

RESULTS. As before, the log odds of a language encoding the visual modality with a unimodal verb were reliably higher than those for nonvisual modalities (log odds: 0.60, $SE = 0.35$; 95% CrI [-0.02, 0.35]), with a very low posterior probability of the effect being negative ($\beta_1 < 0 = 2.8$). There was no reliable difference between ‘hearing’ and ‘touch’, ‘taste’, and ‘smell’ (log odds -0.06, $SE = 0.22$), with the posterior distributions firmly overlapping with zero (95% CrI [-0.50, 0.42]). Languages encoded touch meanings less often with unimodal verbs compared to smell and taste (log odds -0.36, $SE = 0.25$), although this effect was somewhat less robust compared to the PVDB model (95% CrI [-0.86, 0.17]), with a small number of posterior samples above zero ($\beta_1 > 0 = 6.6$). There was also no reliable difference between smell and taste (log odds -0.17, $SE = 0.41$; 95% CrI [-0.96, 0.70]).

In sum, we find the same overall pattern of results in the CLICS³ data as in the PVDB: a strong bias for vision to be lexicalized as a distinct concept, but no tendency for lexicalization rates to mirror the proposed sensory hierarchy.

4.2. CROSSMODAL BIASES IN CLICS³. Next, we constructed a semantic map to visualize the crossmodal associations in the CLICS³ data set. For this, we used the full set of multisense verb forms in CLICS³ (i.e. we included data from all languages in CLICS³ that had one or more multisense verbs). We excluded ten languages also represented in the PVDB, resulting in a final data set of 102 multisense verb forms from ninety-five languages and twenty-eight language families. As in the PVDB, fifteen distinct types of multisense verbs were represented in total, differing from each other with respect to the specific number (range 2–4) and type of sensory meanings encoded (see supplementary materials S7).

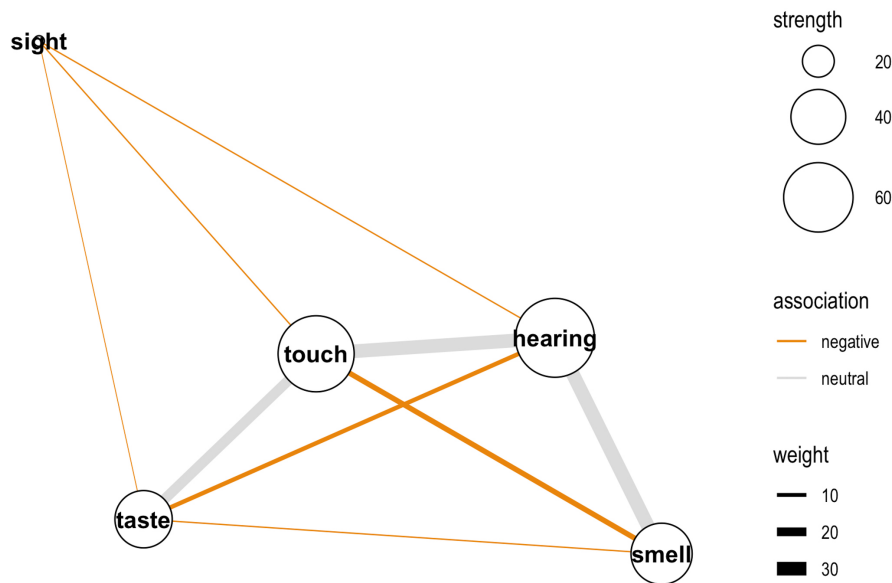


FIGURE 10. Weighted semantic map of crossmodal colexifications in CLICS³. Nodes indicate meanings, and size reflects strength scores. Edges indicate colexifications, thickness reflects crosslinguistic frequency of colexified meanings, and colors indicate whether connected meanings are significantly negatively (orange) or neutrally (gray) associated, according to the pairwise probabilistic coexpression analysis.

Figure 10 shows the same overall pattern as Fig. 8 with respect to node centrality: hearing and touch are the most central nodes (with strength scores of 76 and 71, respectively), taste and smell have lower strength scores (43 and 48, respectively), and sight is the least central node (14). The most salient difference between the figures is the additional edge between sight and taste in CLICS³ (Fig. 10). This edge is supported by three languages (see Table 2). Two of these, Chechen (Nakh-Daghestanian) and Machame (Atlantic Congo), colexify agentive ‘look at’ with ‘taste’ (a colexification not attested in the PVDB). The third, Ossetian (Indo-European) colexifies ‘see’ and ‘taste’. Sight otherwise shows the same set of direct associations as in the PVDB: it is directly linked to hearing and touch, but not smell. The absence of a sight-smell edge underscores the rarity of this crossmodal association.¹⁰

¹⁰ The CLICS³ data in Table 2 contradict descriptions in published sources for at least two cases we are aware of. First, Viberg (1984) writes that Kobon’s verb *nij* (*nōŋ* according to his source) can refer to any sense-modality meaning (though when it refers to modalities other than sight it often carries extra specification; Viberg, p.c.). Second, in Viberg 1984, the Hausa verb *gani* is only listed with the sensory meaning ‘see’, so the inclusion of ‘feel’ might be a glossing error. Viberg (p.c.) informs us that the verb can also mean ‘think’ or ‘have an opinion’ according to his sources, possibly implying an alternative sense of ‘feel’, that is, not tactile perception.

LANGUAGE	FAMILY (AREA)	VERB FORM	SENSORY MEANINGS
Chechen	Nakh-Daghestanian (Eurasia)	həʒə	'look at, taste'
Machame	Atlantic-Congo (Africa)	ilola	
Ossetian	Ossetian (Indo-European)	winin	'see, taste'
Carijona	Cariban (South America)	etae	'see, hear'
Kobon	Nuclear Trans New Guinea (Papunesia)	niŋ-	
Aisi (Musak)	Nuclear Trans New Guinea (Papunesia)	iri-	'see, feel'
Hausa	Afro-Asiatic (Africa)	gani	
Avar (Ansalta)	Nakh-Daghestanian (Eurasia)	bikhize	'see, feel'
Avar (Kunzakh)	Nakh-Daghestanian (Eurasia)	bikhizi	
Udi	Nakh-Daghestanian (Eurasia)	akisun	'see, touch'
Atohwaim	Kayagaric (Papunesia)	saʔap	
Gurdjar	Pama-Nyungan (Australia)	ak	'see, hear, smell'
Mbunga (Kimbunga)	Atlantic-Congo (Africa)	ku	'see, hear, taste, smell'

TABLE 2. Crossmodal colexifications that include visual meanings in CLICS³.

As in Fig. 8, Fig. 10 shows first-order connections between all nonvisual senses, but again the statistical likelihood of colexification is not the same for all pairs. Smell and taste are negatively associated. Only two languages independently colexify smell and taste: Middle High German (*smecken*) and Northern Yukaghir (*morej-*). We find the same relative differences between associates in terms of observed crosslinguistic frequencies, though the association types have shifted in lockstep. The three positively associated pairings in the PVDB model—hearing-touch, hearing-smell, and touch-taste—are neutrally associated in the CLICS³ model; the two neutral associations in the PVDB model—touch-smell and hearing-taste—are negatively associated in the CLICS³ model.

4.3. MACROAREAL TENDENCIES FOUND IN THE PVDB AND CLICS³. The crossmodal semantic maps from the PVDB and CLICS³ together indicate a general bias against the colexification of any modality with sight and against the colexification of taste and smell together. Given that the most robust signal for typological biases holds across different geographical areas (Bickel 2013, Dryer 1989), we asked whether the patterns we find are robust across areas.¹¹ To address this, we combined the two data sets and conducted probabilistic coexpression analyses separately for each macroarea for which there was sufficient data: Africa, Eurasia, Papunesia, and South America. For the remaining two macroareas, Australia and North America, we did not have sufficient data to obtain reliable results (the expected counts for some crossmodal pairs were below five in some instances).

Figure 11 shows heatmaps of the pairwise crossmodal associations for the four macroareas. Green indicates a significant positive association, orange a significant negative association, and gray a neutral (i.e. independent) association. The numbers represent standardized effect sizes, which are calculated as the difference between the expected and observed frequencies of colexification, divided by the number of word forms in the sample.

In all four regions, sight is negatively associated with the nonvisual modalities, and taste and smell are also negatively associated. This consistency across areas strongly indicates a universal bias against the colexification of these sensory meanings. For the remaining crossmodal combinations, we find areal differences. While hearing and touch are positively associated in Eurasia, Papunesia, and South America, they are negatively associated in Africa. By contrast, touch and taste are positively associated in Africa, but negatively associated in Eurasia. Hearing and smell are positively associated in Africa and Papunesia. Hearing and taste show a more consistent pattern across areas: this pairing is negatively associated in Eurasia, and trends negatively in the remaining areas.

¹¹ See Georgakopoulos et al. 2022 for a discussion of areal tendencies based on a smaller subset of languages represented in CLICS².

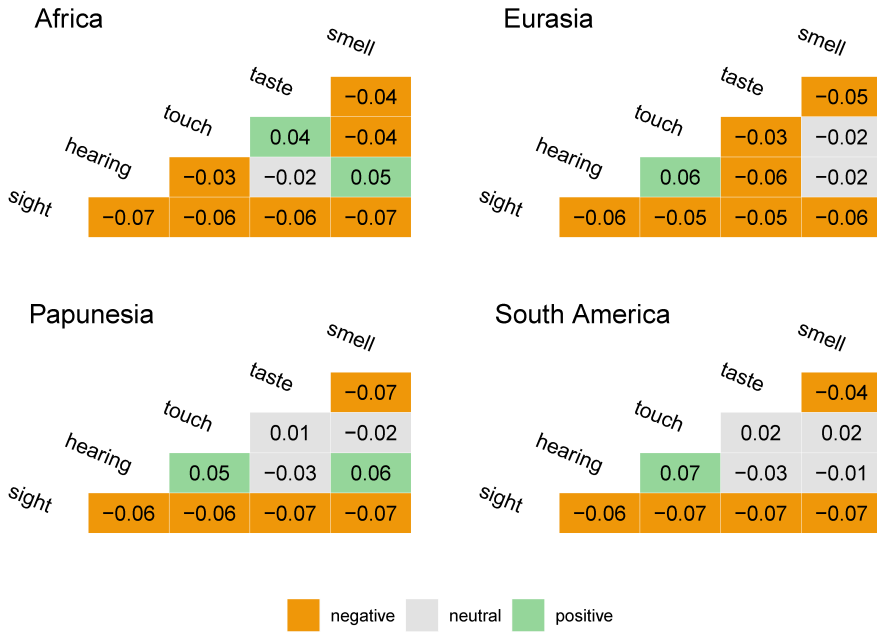


FIGURE 11. Heatmaps showing, for the macroareas of Africa, Eurasia, Papunesia, and South America, the positive (green), negative (orange), and neutral (gray) associations between sensory meanings, as determined by probabilistic coexpression models of the crossmodal colexifications found in the PVDB and CLICS³ combined. Numbers indicate standardized effect sizes.

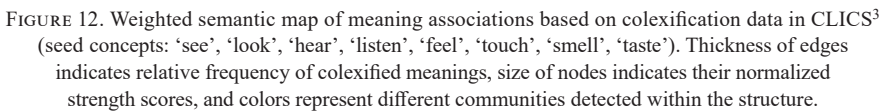
To summarize, we find a strong typological bias against the colexification of vision with other sense modalities, and against the independent colexification of taste and smell. Overlaid on these general biases appear to be distinct areal tendencies for the remaining crossmodal pairs. These areal differences imply that there is no universally preferred crossmodal association. So, while there appear to be strong constraints governing what cannot colexify, the attested colexification of sensory meanings is open to cultural variation.

4.4. TRANSFIELD SEMANTIC CHAINING IN CLICS³. Finally, we checked whether crossmodal colexifications in CLICS³ are the outcome of transfield semantic chaining, following the procedure in §3.4. We include edges only between meanings that had a frequency of three or more in order to take into account the larger set of (closely related) languages and concepts. We also removed all nodes from nonverbal categories (i.e. object and property concepts) to maintain consistency across analyses. For clarity we standardized the names of basic perception concepts, so they matched those in the PVDB in Fig. 9 (e.g. ‘taste (something)’ in CLICS³ became simply ‘taste’). The final data set consisted of 513 languages and 959 word forms.

Once again, five communities were detected, each clustered around a sensory modality (see Figure 12). We find direct links between visual and the nonvisual modalities, albeit at very low frequencies: in contrast to the PVDB, ‘see’ connects directly to ‘hear’, as well as to ‘feel’ and ‘smell’. Inspection of the data reveals that the direct edge between ‘see’ and ‘hear’ is supported by a single data point: the verb *ak* in Gurdjar,¹² which apparently encodes ‘see’, ‘hear’, and ‘smell’, but no transfield meanings. However, this

¹² The language name is spelled Gurdjar in Glottolog and Kurtjar in CLICS³.

[86.4.139.226] Project MUSE (2024-03-17 18:26 GMT) Bodleian Libraries of the University of Oxford



4.5. SUMMARY. The mappings between sensory meanings and perception verbs in CLICS³ largely confirm the findings from the PVDB, lending empirical support for the typological robustness of the observed patterns. We found a strong tendency for vision to be lexically differentiated from nonvisual modalities, but no tendency for lexicalization rates to reflect the proposed sense-modality hierarchy. With respect to crossmodal biases, hearing-touch, hearing-smell, and touch-taste were most frequently

colexified overall, though a joint analysis of the two data sets stratified by macroarea revealed area-specific preferences. Importantly, for all four macroareas tested, taste and smell were negatively associated, indicating a universal bias against the colexification of these meanings. Finally, semantic maps of intrafield and transfield colexifications revealed that, with the exception of vision, crossmodal associations arise primarily via semantic extensions within the perception domain.

5. DISCUSSION. Our analysis of a systematically stratified global sample of 100 languages in the PVDB, replicated with a convenience sample of languages drawn from CLICS³, reveals robust crosslinguistic regularities in the lexical organization of the perception domain. Vision stands apart from other sensory modalities in its encoding, exhibiting a strong bias to be lexicalized with a dedicated verb. For the remaining senses, not all logical pairings are equally likely to be grouped together under a common word. Some crossmodal combinations, in particular hearing-touch, hearing-smell, and touch-taste, are found recurrently across genetically and geographically unrelated languages. By contrast, smell and taste are rarely coexpressed in perception verbs.

Although our results broadly support Viberg's (1984, 2001) original generalization—namely, that across languages there are recurrent asymmetries between sensory modalities in their lexical expression—our study departs from this foundational work in the characterization of those asymmetries. Viberg (2001, 2008, 2015) proposed that the lexicalization of sensory meanings is constrained according to a sense-modality hierarchy, such that vision is the most likely of the senses to be lexically differentiated, followed by hearing and then touch, taste, and smell. While our study confirms vision's unimodal bias, we found no comparable tendency for lexicalization rates of the other sense modalities to follow a hierarchical ordering. This result challenges the view that a fixed, biologically grounded, hierarchy of senses directly influences the lexical expression of perceptual categories.

The fact that certain combinations of sensory meanings regularly share a common label across genetically and geographically diverse languages suggests that the evolution of perception lexicons is shaped in part by convergent conceptual structures across unrelated language communities. This finding contributes to a growing body of literature that has identified crosslinguistic regularities in colexification patterns in other semantic domains (e.g. Koch 2008, Urban 2012, 2021, Wilkins 1996, Xu, Duong, et al. 2020, Youn et al. 2016). We also found that—with the exception of colexifications involving visual meanings—crossmodal associations tend to be first-order, rather than the product of semantic chaining, implying direct conceptual links between sensory meanings.

Our study supports earlier typological generalizations that hearing-smell and touch-taste are more closely associated than hearing-taste and touch-smell, possibly motivated by the presence of bodily contact between perceiver and stimulus (Evans & Wilkins 2000, Viberg 1984). In our larger sample, however, this emerges as a probabilistic tendency across languages and geographical areas, rather than a categorical difference in possible associations.

The most frequently observed crossmodal pairing, hearing-touch, cross-cuts the dimension of contact. This association is, however, consistent with a large literature showing the close perceptual integration between audition and touch (Guest et al. 2002, Jousmäki & Hari 1998, Schürmann et al. 2004, Suzuki, Gyoba, & Sakamoto 2008; see Winter et al. 2017 for a language-oriented discussion of this literature). The perception of both sounds and surface textures has a salient temporal dimension: sound perception is drawn out over time and characterized by internal change, and tactile perception similarly

involves temporally changing forces in contact with the skin (Strik Lievers & Winter 2018, Winter et al. 2017). Possibly reflecting this perceptual closeness, touch adjectives are frequently used to describe sounds in several languages (see Winter 2019 for discussion). Viberg (p.c.) notes that the semantics of ‘feel’ verbs also often covers internal sensations (interoception). Possibly, it is the interoceptive dimension of ‘feel’ that connects it to audition, insofar as hearing might also be conceptualized as experienced internally, as it were, in the ear. These posited similarities require closer empirical scrutiny in the context of the semantic field as a whole.

Crosslinguistically recurrent colexification patterns in the perception domain seem to run contrary to what would be predicted on a cultural-relativism perspective (e.g. Classen 1997, Howes 1991, 2006), according to which ‘sensory perception is a cultural, as well as a physical act’ (Classen 1997:401). In fact, crosslinguistic regularities sit alongside a large space of diverse language-specific colexifications involving various combinations of sensory and nonsensory meanings (Figs. 9 and 12). This recalls Evans’s (2010:528) observation that polysemy is ‘an interesting meeting ground of general cognitive preferences and culture-specific modulations’. We also find different macroareal tendencies. For example, hearing-touch colexifications were overrepresented in Eurasia, South America, and Papunesia but underrepresented in Africa, while in Africa and in Papunesia, hearing-smell colexifications were overrepresented. While these areal tendencies require further comprehensive study, they are consistent with areal patterns of ‘polysemy calquing’ (Koptjevskaja-Tamm & Liljégren 2017) or ‘polysemy copying’ (Heine & Kuteva 2003, 2005) also reported in the perception domain at smaller geographical scales (Evans & Wilkins 2000, Güldemann & Fehn 2017, Raatikainen 2021, Treis 2010, Viberg 1984; see also Georgakopoulos et al. 2022). This geographic variability highlights the important role of cultural diffusion in shaping the typological space of this semantic domain. It also tells us that there is no single universally preferred crossmodal colexification type.

There are, however, universally dispreferred patterns. Our study reveals a robust typological bias against the colexification of vision with all other sense modalities and against the colexification of smell and taste. This goes against widely held generalizations in the literature (e.g. Huomo 2010, Jędrzejowski & Staniewski 2021, Traugott & Dasher 2002, Viberg 1984). While the source of these biases requires further empirical study, we consider it most likely that they have a communicative basis.

The finding that smell and taste are seldom coexpressed in perception verbs, which is unexpected given the previous claims that these two senses are closely associated in words (Viberg 1984, Winter 2019), cannot be due to (a lack of) conceptual resemblance. At the perceptual level, taste and smell are closely aligned (see Winter 2019:249–50 for an overview)—smell interacts with taste in flavor perception (Shepherd 2011). Although the relationship between the two senses is not symmetrical (i.e. orthonasal smell, via sniffing, provides information about nonedible entities, such as fire, gas leaks, etc.), the two senses are highly integrated, and there is overlap in their processing at the neural level (Small & Prescott 2005). There is, moreover, evidence that this close relationship is reflected conceptually and lexically elsewhere. In English, for example, smell and taste are strongly associated in word-association tasks (Nelson, McEvoy, & Schreiber 1998). Crosslinguistic studies of synesthetic adjectives have also shown that taste adjectives are commonly used to describe smells (see Winter 2019 for review).

So, conceptual resemblance cannot explain the universal bias against the colexification of smell and taste. Instead, we suggest that it is precisely because these senses are so closely associated semantically that they are kept apart in basic perception verbs. Brochhagen and

Boleda (2022) propose that when meanings are ‘too similar’, they tend not to colexify, because they are referenced in very similar contexts and hence not easily disambiguated. We suggest that the colexification of smell and taste may be inhibited for this reason. This possibility receives support from sensory norming studies, where speakers rate how much a given property or object is experienced by each of the senses. Importantly, across all languages studied so far, olfactory and gustatory ratings have the strongest positive correlation (Chen et al. 2019, Lynott & Connell 2009, 2013, Lynott et al. 2020, Miklashevsky 2018, Morucci, Bottini, & Crepaldi 2019, Speed & Majid 2017, Vergallito, Petilli, & Marelli 2020). If words evoking olfactory percepts tend to evoke gustatory associations and vice versa, it follows that linguistic context will not reliably disambiguate between an intended smell or taste interpretation of a perception verb that colexifies the two. For example, in an utterance such as *he blinked the sweet mango*, where *blink* is a perception verb that colexifies smell and taste, neither the noun nor its modifier restricts the interpretation of the verb to a single sensory modality. Despite their close conceptual resemblance, the potential for confusability in context therefore inhibits the colexification of smell and taste.

Vision’s unimodal bias can also be explained on communicative grounds—it is predicted straightforwardly by theories of communicative need, which hold that named semantic categories tend to be narrower where a concept is frequently referenced and broader in low-frequency (low-need) parts of the domain (Gibson et al. 2017, Kemp & Regier 2012, Kemp et al. 2018, Zaslavsky et al. 2019a). Across languages, vision verbs have consistently been found to have higher token frequency than nonvisual perception verbs (Holmer 2021, San Roque et al. 2015, Tchantourian & Vamling 2005, Viberg 1993). If perception verb lexicons are oriented to the communicative needs of speakers, it follows that vision would show the greatest tendency toward lexical differentiation within the perception domain. The potential for ambiguity in context also plausibly contributes to vision’s unimodal bias. Sensory norming studies have shown, for all languages studied so far, that object and property words are rated as primarily experienced in the visual modality and that even words that strongly evoke one or more nonvisual modalities tend also to be associated with vision (Chen et al. 2019, Lynott & Connell 2009, 2013, Lynott et al. 2020, Miklashevsky 2018, Morucci et al. 2019, Speed & Majid 2017, Vergallito et al. 2020). Just as with smell and taste, this likely implies that linguistic context cannot offer reliable cues to disambiguate between visual and nonvisual readings of ambiguous perception verbs.

Notably, outside of the domain of perception, vision does frequently colexify (with nonsensory meanings: Figs. 9 and 12; see also Evans & Wilkins 2000, Georgakopoulos et al. 2022, San Roque et al. 2018, Sweetser 1990, Vanhove 2008a). This dual patterning of vision—lexically differentiated within the perception domain but regularly coupled to other meanings outside of it—follows from communicative principles. As Brochhagen and Boleda (2022) observe, because metaphoric meaning extensions (e.g. from vision to cognition) involve meaning pairs that belong to ontologically different domains, colexifying them is unlikely to cause confusion in context.

While the communicative account of these restrictions requires further empirical verification, we note that some qualitative support can be marshalled. Aikhenvald (p.c., cited in Wierzbicka 2010:80) reports that Tariana (Arawak) has a perception verb coexpressing sight and hearing, but, significantly, when the verb is intended in its auditory sense, it is grammatically constrained to select an auditory object of some kind (e.g. words, sounds, language, a voice, etc.). Auditory objects are, importantly, not visually perceivable and hence resolve the sensory interpretation of the perception verb. Similarly, Pawley (2020) reports that Kalam’s general perception verb *ny* is used to express

both seeing and hearing (in addition to other meanings; see §3.3); in order to have a hearing interpretation, it must be used in a two-clause construction, with the object expressed by a ‘sound-making’ verbal clause involving the verb *ag* ‘make a sound, say, utter’. Such examples suggest that in those typologically rare cases where languages do colexify sensory meanings that are frequently confusable in context, they may grammaticalize compensatory strategies (e.g. in the form of selectional restrictions or constructional constraints) to facilitate the recoverability of those meanings. These kinds of disambiguation-driven grammaticalization processes vis-à-vis colexification profiles warrant further typological study. Patterns of partial colexification (involving only partial identity of form), which we have not covered in this study, may similarly be motivated by communicative pressures.

Taken together, we suggest that lexicalization patterns in the perception verb domain can be captured by two domain-general constraints on lexical change: conceptual resemblance (which provides the potential for colexification) and communicative pressures (which may inhibit it). This account is consistent with a growing body of evidence suggesting that languages evolve to efficiently balance between competing cognitive and communicative pressures (Bentz et al. 2017, Carr et al. 2020, Fedzechkina et al. 2012, Gibson et al. 2019, Haspelmath 2021, Kemp & Regier 2012, Kirby et al. 2015, Levshina 2022, Piantadosi et al. 2012, Smith 2020, Zipf 1949) and aligns perception with a diverse set of semantic domains for which this argument has already been developed (Gibson et al. 2017, Kemp & Regier 2012, Regier et al. 2016, Xu, Liu, & Regier 2020, Zaslavsky et al. 2019a).

Could biological rather than communicative factors explain vision’s unimodal bias, as originally proposed by Viberg? In contrast to a fixed hierarchy of senses, there is more empirical support for the notion of visual dominance in human perception (for review see Stokes & Biggs 2014). For example, visual information is often privileged over other senses when sensory information is integrated (e.g. Spence, Parise, & Chen 2012) and, anatomically, vision occupies the largest part of the cortex (Palmer 1999). Perhaps, then, as argued by Viberg, it is vision’s dominance in perception that motivates its bias toward lexical differentiation. The difficulty with this theory is that, on its own, it is unconstrained in the way biological facts are predicted to relate to their linguistic expression. On Viberg’s markedness account, vision, as the dominant sense, is expressed in language in the ‘least marked’ way; lexical differentiation from the other senses is interpreted as an instance of unmarked linguistic behavior. However, nothing in the theory requires markedness to associate with lexicalization patterns in this way. If the facts were to be turned on their heads, such that vision was the MOST likely of the senses to extend its meaning within the perceptual domain, visual dominance in perception could equally well be invoked to account for this fact—either set of behavior could, in principle, be treated as the ‘unmarked’ one (indeed, as we saw in §2.2, the secondary literature has in fact frequently misinterpreted the facts in the opposite direction). On the communicative-need account, by contrast, the direction of the asymmetry is predicted by the theory—informativity is maximized in lexicons where it is needed most.

The communicative-need account does not rule out an indirect role for perceptual factors in shaping perceptual vocabulary, however. Reprising Viberg’s original idea, vision’s dominance in perception plausibly explains why cultures talk most about vision (though see San Roque et al. 2015 and Winter 2019 for other, mutually compatible possibilities), and thus could be considered the distal cause (mediated by frequency of use) of the tendency for visual meanings to be encoded separately in words.

6. CONCLUSION. Despite surface variability, we found strong crosslinguistic biases in patterns of colexification in the perception domain. Vision stood apart in showing a strong crosslinguistic bias to be lexicalized with a dedicated verb. We also found that hearing-touch, touch-taste, and hearing-smell were recurrently colexified across languages, while taste and smell were seldom coexpressed in words. While previous approaches have emphasized the differential dominance of the five senses in explaining lexicalization patterns of perception verbs (Evans & Wilkins 2000, Viberg 1984), we propose that two domain-general constraints—conceptual similarity and communicative pressures—interact to give rise to these patterns.

APPENDIX: LANGUAGES SAMPLED IN THE PVDB		
Macroareas and genetic affiliations are taken from Glottolog 4.8 (Hammarström et al. 2023).		
MACROAREA	LANGUAGES (FAMILIES)	# LGS
Africa	Bokobaru (Mande)	12
	Boro (Ta-Ne Omotic)	
	Central Kanuri (Saharan)	
	Dongola (Nubian)	
	East Taa (Tuu)	
	#Haba (Khoe-Kwadi)	
Australia	Guragone (Maningrida)	8
	Jingulu (Mirndi)	
	Kayardild (Tangkic)	
	Kuku-Yalanji (Pama-Nyungan)	
Eurasia	Abkhaz (Abkhaz-Adyge)	8
	Finnish (Uralic)	
	Halh Mongolian (Mongolic-Khitān)	
	Italian (Indo-European)	
North America	Central Alaskan Yupik (Eskimo-Aleut)	17
	Francisco León Zoque (Mixe-Zoque)	
	Koasati (Muskogean)	
	Mayangna (Misumalpan)	
	Mohave (Cochimi-Yuman)	
	Mutsun (Miwok-Costanoan)	
	Nisga’a (Tsimshian)	
	Northern Haida (Haidan)	
Papunesia	San Mateo del Mar Huave (Huavean)	30
	Ambulas (Ndu)	
	Baruya (Angan)	
	Berik (Tor-Orya)	
	Dera (Senagi)	
	Duna (Isolate)	
	Grass Koiari (Koiarian)	
	Kalam (Nuclear Trans New Guinea)	
	Kaluli (Bosavi)	
	Kaure-Narau (Kaure-Kosare)	
	Kombio (Nuclear Torricelli)	
	Konai (East Strickland)	
	Mailu (Mailuan)	
	Mali (Baining)	
	Maybrat-Karon (Isolate)	
	Mende (Sepik)	

MACROAREA	LANGUAGES (FAMILIES)		# LGS
South America	Arhuaco (Chibchan)	Pumé (Isolate)	25
	Capanahua (Pano-Tacanan)	Qawasqar (Kawesqar)	
	Cofán (Isolate)	Shawi (Cahuapanan)	
	Djeoromixí (Nuclear-Macro-Je)	Shuar (Chicham)	
	Galibi Carib (Cariban)	Tupinambá (Tupian)	
	Guambiano (Barbacoan)	Wapishana (Arawakan)	
	Huallaga Húanuco Quechua (Quechuan)	Warao (Isolate)	
	Kotiria (Tucanoan)	Wichí Lhamtés Vejoz (Matacoan)	
	Mapudungun (Araucanian)	Yagua (Peba-Yagua)	
	Movima (Isolate)	Yanomamö (Yanomamic)	
	Murui Huitoto (Huitotoan)	Yatê (Fulniö)	
	Páez (Isolate)	Yuracaré (Isolate)	
	Paumari (Arawan)		

REFERENCES

- AIKHENVALD, ALEXANDRA Y., and ANNE STORCH. 2013. Linguistic expression of perception and cognition: A typological glimpse. *Perception and cognition in language and culture*, ed. by Alexandra Y. Aikhenvald and Anne Storch, 1–45. Leiden: Brill. DOI: 10.1163/9789004210127_002.
- ARITA, HÉCTOR T. 2016. Species co-occurrence analysis: Pairwise versus matrix-level approaches [Correspondence]. *Global Ecology and Biogeography* 25(11).1397–1400. DOI: 10.1111/geb.12418.
- BENTZ, CHRISTIAN; DIMITRIOS ALIKANIOTIS; MICHAEL CYSOUW; and RAMON FERRER-I-CANCHO. 2017. The entropy of words—Learnability and expressivity across more than 1000 languages. *Entropy* 19(6):275. DOI: 10.3390/e19060275.
- BERLIN, BRENT, and PAUL KAY. 1991. *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.
- BICKEL, BALTHASAR. 2013. Distributional biases in language families. *Language typology and historical contingency: In honor of Johanna Nichols*, ed. by Balthasar Bickel, Lenore A. Grenoble, David A. Peterson, and Alan Timberlake, 415–44. Amsterdam: John Benjamins.
- BLACKINGS, MAIRI JOHN. 2000. *Ma'di-English, English-Ma'di dictionary*. (Languages of the world 25.) Munich: LINCOM Europa.
- BOWERN, CLAIRE. 2010. Historical linguistics in Australia: Trees, networks and their implications. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1559). 3845–54. DOI: 10.1098/rstb.2010.0013.
- BROCHHAGEN, THOMAS, and GEMMA BOLEDA. 2022. When do languages use the same word for different meanings? The Goldilocks principle in colexification. *Cognition* 226:105179. DOI: 10.1016/j.cognition.2022.105179.
- BRUGMAN, CLAUDIA MARLEA. 1988. *The story of over: Polysemy, semantics, and the structure of the lexicon*. (Outstanding dissertations in linguistics.) New York: Garland.
- BÜRKNER, PAUL-CHRISTIAN. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). DOI: 10.18637/jss.v080.i01.
- BÜRKNER, PAUL-CHRISTIAN. 2018. Advanced Bayesian multilevel modeling with the R package brms. *The R Journal* 10(1).395–411. DOI: 10.32614/RJ-2018-017.
- BÜRKNER, PAUL-CHRISTIAN. 2021. Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software* 100(5). DOI: 10.18637/jss.v100.i05.
- BUTAEV, SHAVKAT, and ABBOS IRISQULOV. 2008. *English-Uzbek, Uzbek-English dictionary: 70.000 words and expressions/Inglizcha-O'zbekcha, O'zbekcha-Inglizcha lug'at 70.000 so'z va ibora*. Toshkent: O'zbekiston Respublikasi Fanlar akademiyasi 'Fan' nashriyoti.
- BYBEE, JOAN; REVERE PERKINS; and WILLIAM PAGLIUCA. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- CARR, JON W.; KENNY SMITH; JENNIFER CULBERTSON; and SIMON KIRBY. 2020. Simplicity and informativeness in semantic category systems. *Cognition* 202:104289. DOI: 10.1016/j.cognition.2020.104289.

- CHEN, I-HSUAN; QINGQING ZHAO; YUNFEI LONG; QIN LU; and CHU-REN HUANG. 2019. Mandarin Chinese modality exclusivity norms. *PLOS ONE* 14(2):e0211336. DOI: 10.1371/journal.pone.0211336.
- CLASSEN, CONSTANCE. 1997. Foundations for an anthropology of the senses. *International Social Science Journal* 49(153).401–12. DOI: 10.1111/j.1468-2451.1997.tb00032.x.
- CONWAY, BEVIL R.; SIVALOGESWARAN RATNASINGAM; JULIAN JARA-ETTINGER; RICHARD FUTRELL; and EDWARD GIBSON. 2020. Communication efficiency of color naming across languages provides a new framework for the evolution of color terms. *Cognition* 195:104086. DOI: 10.1016/j.cognition.2019.104086.
- CORBETT, GREVILLE G. 2005. Suppletion in personal pronouns: Theory versus practice, and the place of reproducibility in typology. *Linguistic Typology* 9(1).1–23. DOI: 10.1515/lity.2005.9.1.1.
- CROFT, WILLIAM. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- CROFT, WILLIAM. 2002. *Typology and universals*. 2nd edn. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511840579.
- CROFT, WILLIAM. 2022. On two mathematical representations for ‘semantic maps’. *Zeitschrift für Sprachwissenschaft* 41(1).67–87. DOI: 10.1515/zfs-2021-2040.
- CSÁRDI, GÁBOR, and TAMÁS NEPUSZ. 2006. The igraph software package for complex network research. *Interjournal, Complex Systems*, 1695. Online: <https://igraph.org/>.
- CYFFER, NORBERT, and JOHN HUTCHINSON. 1990. *Dictionary of the Kanuri language*. Berlin: De Gruyter. DOI: 10.1515/9783110869231.
- DRYER, MATTHEW S. 1989. Large linguistic areas and language sampling. *Studies in Language* 13(2).257–92. DOI: 10.1075/sl.13.2.03dry.
- EVANS, NICHOLAS. 1988. Arguments for Pama-Nyungan as a genetic subgroup, with particular reference to initial laminalization. *Aboriginal linguistics 1*, ed. by Nicholas Evans and Steve Johnson, 91–110. Armidale: University of New England, Department of Linguistics. Online: <http://hdl.handle.net/1885/10440>.
- EVANS, NICHOLAS. 2007. Warramurrungunji undone: Australian languages in the 51st millennium. *Language diversity endangered*, ed. by Matthias Brenzinger, 342–73. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110905694-017.
- EVANS, NICHOLAS. 2010. Semantic typology. *The Oxford handbook of linguistic typology*, ed. by Jae Jung Song, 504–33. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199281251.013.0024.
- EVANS, NICHOLAS, and DAVID WILKINS. 2000. In the mind’s ear: The semantic extensions of perception verbs in Australian languages. *Language* 76(3).546–92. DOI: 10.2307/417135.
- FEDZECHKINA, MARYIA; T. FLORIAN JAEGER; and ELISSA L. NEWPORT. 2012. Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences* 109(44).17897–902. DOI: 10.1073/pnas.1215776109.
- FLOYD, SIMEON; LILA SAN ROQUE; and ASIFA MAJID. 2018. Smell is coded in grammar and frequent in discourse: Cha’palaa olfactory language in cross-linguistic perspective. *Journal of Linguistic Anthropology* 28(2).175–96. DOI: 10.1111/jola.12190.
- FRANÇOIS, ALEXANDRE. 2008. Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In Vanhove 2008b, 163–215.
- GAST, VOLKER, and MARIA KOPTJEVSKAJA-TAMM. 2022. Patterns of persistence and diffusibility in the European lexicon. *Linguistic Typology* 26(2).403–38. DOI: 10.1515/lingty-2021-2086.
- GEERAERTS, DIRK. 1997. *Diachronic prototype semantics: A contribution to historical lexicology*. (Oxford studies in lexicography and lexicology.) Oxford: Oxford University Press.
- GENTNER, DEDRE, and MELISSA BOWERMAN. 2009. Why some spatial semantic categories are harder to learn than others: The typological prevalence hypothesis. *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*, ed. by Jiansheng Guo, Elena Lieven, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura, and Seyda Ozcaliskan, 465–80. New York: Psychology Press.
- GEORGAKOPOULOS, THANASIS; EITAN GROSSMAN; DMITRY NIKOLAEV; and STÉPHANE POLIS. 2022. Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain. *Linguistic Typology* 26(2).439–87. DOI: 10.1515/lingty-2021-2088.

- GEORGAKOPOULOS, THANASIS, and STÉPHANE POLIS. 2018. The semantic map model: State of the art and future avenues for linguistic research. *Language and Linguistics Compass* 12(2):e12270. DOI: 10.1111/lnc3.12270.
- GEURTS, KATHRYN LINN. 2002. *Culture and the senses: Bodily ways of knowing in an African community*. (Ethnographic studies in subjectivity 3.) Berkeley: University of California Press.
- GIBSON, EDWARD; RICHARD FUTRELL; JULIAN JARA-ETTINGER; KYLE MAHOWALD; LEON BERGEN; SIVALOGESWARAN RATNASINGAM; MITCHELL GIBSON; STEVEN T. PIANTADOSI; and BEVIL R. CONWAY. 2017. Color naming across languages reflects color use. *Proceedings of the National Academy of Sciences* 114(40):10785–90. DOI: 10.1073/pnas.1619666114.
- GIBSON, EDWARD; RICHARD FUTRELL; STEVEN P. PIANTADOSI; ISABELLE DAUTRICHE; KYLE MAHOWALD; LEON BERGEN; and ROGER LEVY. 2019. How efficiency shapes human language. *Trends in Cognitive Sciences* 23(5):389–407. DOI: 10.1016/j.tics.2019.02.003.
- GREENBERG, JOSEPH HAROLD. 1966. *Language universals: With special reference to feature hierarchies*. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110899771.
- GRIFFITH, DANIEL M.; JOSEPH A. VEECH; and CHARLES J. MARSH. 2016. cooccur: Probabilistic species co-occurrence analysis in R. *Journal of Statistical Software, Code Snippets* 69(2). DOI: 10.18637/jss.v069.c02.
- GUEST, STEVE; CAROLINE CATMUR; DONNA LLOYD; and CHARLES SPENCE. 2002. Audiotactile interactions in roughness perception. *Experimental Brain Research* 146(2):161–71. DOI: 10.1007/s00221-002-1164-z.
- GÜLDEMANN, TOM, and ANNE-MARIA FEHN. 2017. The Kalahari Basin Area as a ‘sprachbund’ before the Bantu Expansion. *The Cambridge handbook of areal linguistics*, ed. by Raymond Hickey, 500–526. Cambridge: Cambridge University Press. DOI: 10.1017/9781107279872.019.
- GYÖRI, GÁBOR. 2002. Semantic change and cognition. *Cognitive Linguistics* 13(2):123–66. DOI: 10.1515/cogl.2002.012.
- HAMID, ABDUL. 1998. *Ladakhi-English-Urdu dictionary, with an English-Ladakhi index*. Leh: Melong.
- HAMMARSTRÖM, HARALD, and MARK DONOHUE. 2014. Some principles on the use of macro-areas in typological comparison. *Language Dynamics and Change* 4(1):167–87. DOI: 10.1163/22105832-00401001.
- HAMMARSTRÖM, HARALD; ROBERT FORKEL; MARTIN HASPELMATH; and SEBASTIAN BANK. 2023. Glottolog 4.8. Leipzig: Max Planck Institute for Evolutionary Anthropology. DOI: 10.5281/zenodo.8131084.
- HASPELMATH, MARTIN. 1997. *Indefinite pronouns*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198235606.001.0001.
- HASPELMATH, MARTIN. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. *The new psychology of language: Cognitive and functional approaches to language structure*, vol. 2, ed. by Michael Tomasello, 211–43. Mahwah, NJ: Lawrence Erlbaum.
- HASPELMATH, MARTIN. 2021. Explaining grammatical coding asymmetries: Form–frequency correspondences and predictability. *Journal of Linguistics* 57(3):605–33. DOI: 10.1017/S0022226720000535.
- HASPELMATH, MARTIN, and SVEN SIEGMUND. 2006. Simulating the replication of some of Greenberg’s word order generalizations. *Linguistic Typology* 10(1):74–82. DOI: 10.1515/LINGTY.2006.004.
- HEINE, BERND, and TANIA KUTEVA. 2003. On contact-induced grammaticalization. *Studies in Language* 27(3):529–72. DOI: 10.1075/sl.27.3.04hei.
- HEINE, BERND, and TANIA KUTEVA. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511614132.
- HERSHBERGER, HENRY D., and RUTH HERSHBERGER. 1982. *Kuku-Yalanji dictionary*. (Work Papers of SIL-AAB 7.) Darwin: Summer Institute of Linguistics, Australian Aborigines Branch.
- HILL, JANE H. 1988. Language, culture, and world-view. *Linguistics: The Cambridge survey*, vol. 4. *Language: The socio-cultural context*, ed. by Frederick J. Newmeyer, 14–36. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620577.003.

- HOLMER, SONJA. 2021. The language of vision in four Aslian speech communities: An introductory investigation of basic vision verbs. Lund: Lund University B.A. thesis. Online: <http://lup.lub.lu.se/student-papers/record/9041467>.
- HOPPER, PAUL J., and ELIZABETH CLOSS TRAUGOTT. 2003. *Grammaticalization*. 2nd edn. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139165525.
- HOWES, DAVID (ed.) 1991. *The varieties of sensory experience: A sourcebook in the anthropology of the senses*. (Anthropological horizons.) Toronto: University of Toronto Press.
- HOWES, DAVID. 2006. Charting the sensorial revolution. *The Senses and Society* 1(1).113–28. DOI: 10.2752/174589206778055673.
- HUUMO, TUOMAS. 2010. Is perception a directional relationship? On directionality and its motivation in Finnish expressions of sensory perception. *Linguistics* 48(1).49–97. DOI: 10.1515/ling.2010.002.
- JACKSON, JOSHUA CONRAD; JOSEPH WATTS; TEAGUE R. HENRY; JOHANN-MATTIS LIST; ROBERT FORKEL; PETER J. MUCHA; SIMON J. GREENHILL; RUSSELL D. GRAY; and KRISTEN A. LINDQUIST. 2019. Emotion semantics show both cultural variation and universal structure. *Science* 366(6472).1517–22. DOI: 10.1126/science.aaw8160.
- JĘDRZEJOWSKI, ŁUKASZ, and PRZEMYSŁAW STANIEWSKI. 2021. Rendering what the nose perceives: An introduction. *The linguistics of olfaction: Typological and diachronic approaches to synchronic diversity*, ed. by Łukasz Jędrzejowski and Przemysław Staniewski, 1–34. Amsterdam: John Benjamins.
- JOUSMÄKI, V., and R. HARI. 1998. Parchment-skin illusion: Sound-biased touch. *Current Biology* 8(6).R190–R191. DOI: 10.1016/S0960-9822(98)70120-4.
- JURAFSKY, DANIEL. 1996. Universal tendencies in the semantics of the diminutive. *Language* 72(3).533–78. DOI: 10.2307/416278.
- KARJUS, ANDRES; RICHARD A. BLYTHE; SIMON KIRBY; TIANYU WANG; and KENNY SMITH. 2021. Conceptual similarity and communicative need shape colexification: An experimental study. *Cognitive Science* 45(9):e13035. DOI: 10.1111/cogs.13035.
- KEMP, CHARLES, and TERRY REGIER. 2012. Kinship categories across languages reflect general communicative principles. *Science* 336(6084).1049–54. DOI: 10.1126/science.1218811.
- KEMP, CHARLES; YANG XU; and TERRY REGIER. 2018. Semantic typology and efficient communication. *Annual Review of Linguistics* 4(1).109–28. DOI: 10.1146/annurev-linguistics-011817-045406.
- KIRBY, SIMON; MONICA TAMARIZ; HANNAH CORNISH; and KENNY SMITH. 2015. Compression and communication in the cultural evolution of linguistic structure. *Cognition* 141.87–102. DOI: 10.1016/j.cognition.2015.03.016.
- KOCH, PETER. 2008. Cognitive onomasiology and lexical change: Around the eye. In Vanhove 2008b, 107–37.
- KOPTJEVSKAJA-TAMM, MARIA, and HENRIK LILJEGREN. 2017. Semantic patterns from an areal perspective. *The Cambridge handbook of areal linguistics*, ed. by Raymond Hickey, 204–36. Cambridge: Cambridge University Press. DOI: 10.1017/9781107279872.009.
- KRISHNA, P. PHANI; S. ARULMOZI; and RAMESH KUMAR MISHRA. 2022. Do you see and hear more? A study on Telugu perception verbs. *Journal of Psycholinguistic Research* 51(3).473–84. DOI: 10.1007/s10936-021-09827-7.
- LAKOFF, GEORGE. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- LEMOINE, NATHAN P. 2019. Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos* 128(7).912–28. DOI: 10.1111/oik.05985.
- LEVSHINA, NATALIA. 2022. *Communicative efficiency: Language structure and use*. Cambridge: Cambridge University Press. DOI: 10.1017/9781108887809.
- LIST, JOHANN-MATTIS. 2023. Inference of partial colexifications from multilingual word-lists. *Frontiers in Psychology* 14:1156540. DOI: 10.3389/fpsyg.2023.1156540.
- LIST, JOHANN-MATTIS; ANSELM TERHALLE; and MATTHIAS URBAN. 2013. Using network approaches to enhance the analysis of cross-linguistic polysemies. *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)—Short papers*, 347–53. Online: <https://aclanthology.org/W13-0208>.

- LUO, YONGXIAN (ed.) 1999. *A dictionary of Dehong, Southwest China*. (Pacific linguistics C-145.) Canberra: Australian National University. Online: <https://openresearch-repository.anu.edu.au/bitstream/1885/146619/1/PL-C145.pdf>.
- LYNOTT, DERMOT, and LOUISE CONNELL. 2009. Modality exclusivity norms for 423 object properties. *Behavior Research Methods* 41(2).558–64. DOI: 10.3758/BRM.41.2.558.
- LYNOTT, DERMOT, and LOUISE CONNELL. 2013. Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods* 45(2).516–26. DOI: 10.3758/s13428-012-0267-0.
- LYNOTT, DERMOT; LOUISE CONNELL; MARC BRYLSBAERT; JAMES BRAND; and JAMES CARNEY. 2020. The Lancaster Sensorimotor Norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods* 52(3).1271–91. DOI: 10.3758/s13428-019-01316-z.
- MAJID, ASIFA; SEÁN G. ROBERTS; LUDY CILISSEN; KAREN EMMOREY; BRENDA NICODEMUS; LUCINDA O'GRADY; BENCIE WOLL; BARBARA LELAN; ET AL. 2018. Differential coding of perception in the world's languages. *Proceedings of the National Academy of Sciences* 115(45).11369–76. DOI: 10.1073/pnas.1720419115.
- MCELREATH, RICHARD. 2020. *Statistical rethinking: A Bayesian course with examples in R and Stan*. New York: Chapman and Hall/CRC. DOI: 10.1201/9780429029608.
- MIESTAMO, MATTI; DIK BAKKER; and ANTTI ARPPE. 2016. Sampling for variety. *Linguistic Typology* 20(2).233–96. DOI: 10.1515/lingty-2016-0006.
- MIKLASHEVSKY, ALEX. 2018. Perceptual experience norms for 506 Russian nouns: Modality rating, spatial localization, manipulability, imageability and other variables. *Journal of Psycholinguistic Research* 47(3).641–61. DOI: 10.1007/s10936-017-9548-1.
- MORUCCI, PIERMATTEO; ROBERTO BOTTINI; and DAVIDE CREPALDI. 2019. Augmented modality exclusivity norms for concrete and abstract Italian property words. *Journal of Cognition* 2(1).42. DOI: 10.5334/joc.88.
- MÜNCH, ALLA, and JOHANNES DELLERT. 2015. Evaluating the potential of a large-scale polysemy network as a model of plausible semantic shifts. Paper presented at the 6th Conference on Quantitative Investigations in Theoretical Linguistics (QITL-6), November 4–6, Tübingen.
- NELSON, DOUGLAS L.; CATHY L. McEVROY; and THOMAS A. SCHREIBER. 1998. The University of South Florida word association, rhyme, and word fragment norms. Online: <http://w3.usf.edu/FreeAssociation/>.
- NORCLIFFE, ELISABETH, and ASIFA MAJID. 2023. Word formation patterns in the perception domain: A typological study of cross-modal semantic associations. Oxford: University of Oxford, ms.
- O'MEARA, CAROLYN; LAURA J. SPEED; LILA SAN ROQUE; and ASIFA MAJID. 2019. Perception metaphors: A view from diversity. *Perception metaphors*, ed. by Laura J. Speed, Carolyn O'Meara, Lila San Roque, and Asifa Majid, 1–16. Amsterdam: John Benjamins.
- ONG, WALTER J. 1967. *The presence of the word: Some prolegomena for cultural and religious history*. New Haven, CT: Yale University Press.
- PALMER, STEPHEN E. 1999. *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.
- PAWLEY, ANDREW. 2020. The depiction of sensing events in English and Kalam. *Meaning, life and culture: In conversation with Anna Wierzbicka*, ed. by Helen Bromhead and Zhengdao Ye, 381–402. Canberra: Australian National University Press. Online: <https://www.jstor.org/stable/j.ctv1d5nm0d.26>.
- PEDERSEN, THOMAS LIN. 2021. ggraph: An implementation of grammar of graphics for graphs and networks. R package version 2.0.5. Online: <https://CRAN.R-project.org/package=ggraph>.
- PIANTADOSI, STEVEN T.; HARRY TILY; and EDWARD GIBSON. 2012. The communicative function of ambiguity in language. *Cognition* 122(3).280–91. DOI: 10.1016/j.cognition.2011.10.004.
- R CORE TEAM. 2022. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online: <https://www.r-project.org/>.
- RAATIKAINEN, OONA. 2021. Colexification patterns of perception verbs in the Circum-Baltic area. Helsinki: University of Helsinki M.A. thesis. DOI: 10138/337473.

- RAMIRO, CHRISTIAN; MAHESH SRINIVASAN; BARBARA C. MALT; and YANG XU. 2018. Algorithms in the historical emergence of word senses. *Proceedings of the National Academy of Sciences* 115(10).2323–28. DOI: 10.1073/pnas.1714730115.
- REGIER, TERRY; ALEXANDRA CARSTENSEN; and CHARLES KEMP. 2016. Languages support efficient communication about the environment: Words for snow revisited. *PLOS ONE* 11(4):e0151138. DOI: 10.1371/journal.pone.0151138.
- REGIER, TERRY; PAUL KAY; and NAVEEN KHETARPAL. 2007. Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences* 104(4).1436–41. DOI: 10.1073/pnas.0610341104.
- REGIER, TERRY; NAVEEN KHETARPAL; and ASIFA MAJID. 2013. Inferring semantic maps. *Linguistic Typology* 17(1).89–105. DOI: 10.1515/lity-2013-0003.
- REILLY, JAMIE; MAURICE FLURIE; and JONATHAN E. PELLE. 2020. The English lexicon mirrors functional brain activation for a sensory hierarchy dominated by vision and audition: Point-counterpoint. *Journal of Neurolinguistics* 55:100895. DOI: 10.1016/j.jneuroling.2020.100895.
- RIEMER, NICK. 2010. *Introducing semantics*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511808883.
- ROBERSON, DEBI; IAN DAVIES; and JULES DAVIDOFF. 2000. Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General* 129(3).369–98. DOI: 10.1037/0096-3445.129.3.369.
- RZYMSKI, CHRISTOPH; TIAGO TRESOLDI; SIMON J. GREENHILL; MEI-SHIN WU; NATHANAE E. SCHWEIKHARD; MARIA KOPTJEVSKAJA-TAMM; VOLKER GAST; TIMOTHEUS A. BODT; ET AL. 2020. The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies. *Scientific Data* 7(1):13. DOI: 10.1038/s41597-019-0341-x.
- SAN ROQUE, LILA; KOBIN H. KENDRICK; ELISABETH NORCLIFFE; PENELOPE BROWN; REBECCA DEFINA; MARK DINGEMANSE; TYKO DIRKSMEYER; N. J. ENFIELD; ET AL. 2015. Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics* 26(1).31–60. DOI: 10.1515/cog-2014-0089.
- SAN ROQUE, LILA; KOBIN H. KENDRICK; ELISABETH NORCLIFFE; and ASIFA MAJID. 2018. Universal meaning extensions of perception verbs are grounded in interaction. *Cognitive Linguistics* 29(3).371–406. DOI: 10.1515/cog-2017-0034.
- SAXTON, DEAN; LUCILLE SAXTON; SUSIE ENOS; and R. L. CHERRY. 1983. *Dictionary: Tohono O'odham/Pima to English, English to Tohono O'odham/Pima*. Tucson: University of Arizona Press.
- SCHAPPER, ANTOINETTE. 2019. The ethno-linguistic relationship between smelling and kissing: A Southeast Asian case study. *Oceanic Linguistics* 58(1).92–109. DOI: 10.1353/ol.2019.0004.
- SCHÜRMANN, MARTIN; GINA CAETANO; VEIKKO JOUSMÄKI; and RIITTA HARI. 2004. Hands help hearing: Facilitatory audiotactile interaction at low sound-intensity levels. *The Journal of the Acoustical Society of America* 115(2).830–32. DOI: 10.1121/1.1639909.
- SHEPHERD, GORDON M. 2011. *Neurogastronomy: How the brain creates flavor and why it matters*. New York: Columbia University Press.
- SMALL, DANA M., and JOHN PRESCOTT. 2005. Odor/taste integration and the perception of flavor. *Experimental Brain Research* 166(3–4).345–57. DOI: 10.1007/s00221-005-2376-9.
- SMITH, KENNY. 2020. How culture and biology interact to shape language and the language faculty. *Topics in Cognitive Science* 12(2).690–712. DOI: 10.1111/tops.12377.
- SONG, JAE JUNG. 2007. What or where can we do better? Some personal reflections on (the tenth anniversary of) Linguistic Typology. *Linguistic Typology* 11(1).5–22. DOI: 10.1515/LINGTY.2007.002.
- SPEED, LAURA J., and ASIFA MAJID. 2017. Dutch modality exclusivity norms: Simulating perceptual modality in space. *Behavior Research Methods* 49(6).2204–18. DOI: 10.3758/s13428-017-0852-3.
- SPENCE, CHARLES; CESARE PARISE; and YI-CHUAN CHEN. 2012. The Colavita visual dominance effect. *The neural bases of multisensory processes*, ed. by Micah M. Murray and Mark T. Wallace, 529–56. Boca Raton, FL: CRC Press/Taylor & Francis. Online: <https://www.ncbi.nlm.nih.gov/books/NBK92851/>.

- SRINIVASAN, MAHESH, and HUGH RABAGLIATI. 2015. How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. *Lingua* 157.124–52. DOI: 10.1016/j.lingua.2014.12.004.
- STEBBINS, TONYA N., and JULIUS TAYUL. 2012. *Mali (Baining) dictionary: Mali-Baining Amēthamon Angētha Thēvaik*. (Asia-Pacific Linguistics open access monographs 001.) Canberra: Australian National University. Online: <http://hdl.handle.net/1885/9273>.
- STOKES, DUSTIN; and STEPHEN BIGGS. 2014. The dominance of the visual. *Perception and its modalities*, ed. by Dustin Stokes, Mohan Matthen, and Stephen Biggs, 350–78. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199832798.003.0015.
- STRIK LIEVERS, FRANCESCA, and BODO WINTER. 2018. Sensory language across lexical categories. *Lingua* 204.45–61. DOI: 10.1016/j.lingua.2017.11.002.
- SUZUKI, YUIKA; JIRO GYOKA; and SHUICHI SAKAMOTO. 2008. Selective effects of auditory stimuli on tactile roughness perception. *Brain Research* 1242.87–94. DOI: 10.1016/j.brainres.2008.06.104.
- SWEETSER, EVE. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. (Cambridge studies in linguistics 54.) Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620904.
- TCHANTOURIAN, REVAZ, and KARINA VAMLING. 2005. Basic verb frequency in Megrelian. *Lund Working Papers in Linguistics* 51.199–207. Online: <https://journals.lub.lu.se/LWPL/article/view/2348>.
- TEN RAA, ERIC. 2012. *A dictionary of Sandawe: The lexicon and culture of a Khoesan people of Tanzania*. Ed. by Christopher Ehret and Patricia Ehret. (Research in Khoisan studies/Quellen zur Khoisan-Forschung 27.) Cologne: Rüdiger Köppe.
- TRAILL, ANTHONY. 2018. *A trilingual !Xóǝ dictionary: !Xóǝ-English-Setswana*. Cologne: Rüdiger Köppe.
- TRAUGOTT, ELIZABETH CLOSS, and RICHARD B. DASHER. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486500.
- TREIS, YVONNE. 2010. Perception verbs and taste adjectives in Kambaata and beyond. *Perception of the invisible: Religion, historical semantics and the role of perceptive verbs*, ed. by Anne Storch, 313–46. Cologne: Rüdiger Köppe.
- TWOMEY, COLIN R.; GARETH ROBERTS; DAVID H. BRAINARD; and JOSHUA B. PLOTKIN. 2021. What we talk about when we talk about colors. *Proceedings of the National Academy of Sciences* 118(39):e2109237118. DOI: 10.1073/pnas.2109237118.
- URBAN, MATTHIAS. 2012. *Analyzability and semantic associations in referring expressions: A study in comparative lexicology*. Leiden: Leiden University dissertation. Online: <https://hdl.handle.net/1887/19940>.
- URBAN, MATTHIAS. 2021. Red, black, and white hearts: ‘Heart’, ‘liver’, and ‘lungs’ in typological and areal perspective. *Linguistic Typology* 26(2).349–74. DOI: 10.1515/lingty-2021-2081.
- VANHOVE, MARTINE. 2008a. Semantic associations between sensory modalities, prehension and mental perceptions. In Vanhove 2008b, 341–70.
- VANHOVE, MARTINE (ed.) 2008b. *From polysemy to semantic change: Towards a typology of lexical semantic associations*. Amsterdam: John Benjamins.
- VASISHTH, SHRAVAN; BRUNO NICENBOIM; MARY E. BECKMAN; FANGFANG LI; and EUN JONG KONG. 2018. Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics* 71.147–61. DOI: 10.1016/j.wocn.2018.07.008.
- VERGALLITO, ALESSANDRA; MARCO ALESSANDRO PETILLI; and MARCO MARELLI. 2020. Perceptual modality norms for 1,121 Italian words: A comparison with concreteness and imageability scores and an analysis of their impact in word processing tasks. *Behavior Research Methods* 52(4).1599–1616. DOI: 10.3758/s13428-019-01337-8.
- VIBERG, ÅKE. 1984. The verbs of perception: A typological study. *Explanations for language universals*, ed. by Brian Butterworth, Bernard Comrie, and Östen Dahl, 123–62. DOI: 10.1515/9783110868555.123.
- VIBERG, ÅKE. 1993. Crosslinguistic perspectives on lexical organization and lexical progression. *Progression & regression in language: Sociocultural, neuropsychological, & linguistic perspectives*, ed. by Kenneth Hyldenstam and Åke Viberg, 340–85. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511627781.015.

- VIBERG, ÅKE. 2001. Verbs of perception. *Language typology and language universals: An international handbook*, vol. 2, ed. by Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible, 1294–1309. Berlin: De Gruyter. DOI: 10.1515/9783110194265-030.
- VIBERG, ÅKE. 2008. Swedish verbs of perception from a typological and contrastive perspective. *Languages and cultures in contrast and comparison*, ed. by María de los Angeles Gómez González, J. Lachlan Mackenzie, and Elsa M. González Álvarez, 123–72. Amsterdam: John Benjamins.
- VIBERG, ÅKE. 2015. Sensation, perception and cognition: Swedish in a typological-contrastive perspective. *Functions of Language* 22(1).96–131. DOI: 10.1075/fol.22.1.05vib.
- VIBERG, ÅKE. 2019. Phenomenon-based perception verbs in Swedish from a typological and contrastive perspective. *Syntaxe et sémantique* 20(1).17–48. Online: <https://www.cairn.info/revue-syntaxe-et-semantique-2019-1-page-17.htm>.
- VON HOUWALD, GÖTZ DIETER. 1980. *Diccionario español-sumu, sumu-español*. Managua: Ministerio de Educación.
- WÄLCHLI, BERNHARD. 2016. Non-specific, specific and obscured perception verbs in Baltic languages. *Baltic Linguistics* 7.53–135. DOI: 10.32798/bl.384.
- WICKHAM, HADLEY; MARA AVERICK; JENNIFER BRYAN; WINSTON CHANG; LUCY D'AGOSTINO MCGOWAN; ROMAIN FRANÇOIS; GARRETT GROLEMUND; ALEX HAYES; ET AL. 2019. Welcome to the tidyverse. *Journal of Open Source Software* 4(43):1686. DOI: 10.21105/joss.01686.
- WIERZBICKA, ANNA. 2010. *Experience, evidence, and sense: The hidden cultural legacy of English*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780195368000.001.0001.
- WILKINS, DAVID P. 1996. Natural tendencies of semantic change and the search for cognates. *The comparative method reviewed: Regularity and irregularity in language change*, ed. by Mark Durie and Malcolm Ross, 264–304. Oxford: Oxford University Press. DOI: 10.1093/oso/9780195066074.003.0010.
- WINTER, BODO. 2019. *Sensory linguistics: Language, perception and metaphor*. (Converging evidence in language and communication research.) Amsterdam: John Benjamins. DOI: 10.1075/celcr.20.
- WINTER, BODO; MARCUS PERLMAN; and ASIFA MAJID. 2018. Vision dominates in perceptual language: English sensory vocabulary is optimized for usage. *Cognition* 179.213–20. DOI: 10.1016/j.cognition.2018.05.008.
- WINTER, BODO; MARCUS PERLMAN; LYNN K. PERRY; and GARY LUPYAN. 2017. Which words are most iconic? Iconicity in English sensory words. *Interaction Studies* 18(3).443–64. DOI: 10.1075/is.18.3.07win.
- XU, YANG; KHANG DUONG; BARBARA C. MALT; SERENA JIANG; and MAHESH SRINIVASAN. 2020. Conceptual relations predict colexification across languages. *Cognition* 201: 104280. DOI: 10.1016/j.cognition.2020.104280.
- XU, YANG; EMMY LIU; and TERRY REGIER. 2020. Numeral systems across languages support efficient communication: From approximate numerosity to recursion. *Open Mind* 4.57–70. DOI: 10.1162/opmi_a_00034.
- YANG, ZHAO; RENÉ ALGESHEIMER; and CLAUDIO J. TESSONE. 2016. A comparative analysis of community detection algorithms on artificial networks. *Scientific Reports* 6(1):30750. DOI: 10.1038/srep30750.
- YOUN, HYEJIN; LOGAN SUTTON; ERIC SMITH; CRISTOPHER MOORE; JON F. WILKINS; IAN MADDIESON; WILLIAM CROFT; and TANMOY BHATTACHARYA. 2016. On the universal structure of human lexical semantics. *Proceedings of the National Academy of Sciences* 113(7).1766–71. DOI: 10.1073/pnas.1520752113.
- ZASLAVSKY, NOGA; CHARLES KEMP; NAFTALI TISHBY; and TERRY REGIER. 2019a. Color naming reflects both perceptual structure and communicative need. *Topics in Cognitive Science* 11(1).207–19. DOI: 10.1111/tops.12395.
- ZASLAVSKY, NOGA; CHARLES KEMP; NAFTALI TISHBY; and TERRY REGIER. 2019b. Communicative need in colour naming. *Cognitive Neuropsychology* 37(5–6).312–24. DOI: 10.1080/02643294.2019.1604502.

- ZESHAN, ULRIKE, and NICK PALFREYMAN. 2019. Sensory perception metaphors in sign languages. *Perception metaphors*, ed. by Carolyn O'Meara, Lila San Roque, Asifa Majid, and Laura J. Speed, 275–301. Amsterdam: John Benjamins.
- ZIPF, GEORGE KINGSLEY. 1949. *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley.
- ZWARTS, JOOST. 2010. Semantic map geometry: Two approaches. *Linguistic Discovery* 8(1).377–95. DOI: 10.1349/PS1.1537-0852.A.357.

Norcliffe
Department of Experimental Psychology
Anna Watts Building
Radcliffe Observatory Quarter
Woodstock Road
Oxford, OX2 6GG, UK
[elisabeth.norcliffe@psy.ox.ac.uk]
[asifa.majid@psy.ox.ac.uk]

[Received 12 October 2021;
revision invited 11 May 2022;
revision received 19 December 2022;
revision invited 20 May 2023;
revision received 20 July 2023;
accepted pending revisions 13 August 2023;
revision received 11 September 2023;
accepted 12 September 2023]