The Origins of Inflectional Classes

Xavier Bach
Lincoln College

A thesis submitted for the degree of
Doctor of Philosophy
2018

The Origins of Inflectional Classes.
Xavier C.A. Bach. Lincoln College.
A thesis submitted for the degree of Doctor of Philosophy. Trinity Term 2018.

Abstract

Inflectional classes are classes of lexemes which share a content paradigm (they inflect for the same set of feature values) but differ in their form paradigm (the exponents for these feature values are at least partially different). Irregular lexemes will be viewed as classes with a very low number of members, which do not differ theoretically from classes with larger membership. Contrary to recent developments in morphological theory, inflectional classes are not viewed as classes of stems but classes of lexemes.

Inflectional classes have long been known to arise from sound change applied to inflectional paradigms, creating allomorphy. Other sources identified in the literature include grammaticalization and reanalysis, which have been identified in connection with Germanic languages.

This thesis investigates the origins of inflectional classes from a crosslinguistic perspective, including data on the Romance languages, Pama-Nyungan languages, Austronesian languages, various families of non-Austronesian languages from New Guinea, Niger-Congo languages, and a number of other families. It confirms that sound change, grammaticalization, and reanalysis are major sources for inflectional classes.

It shows that sound change is major source of inflectional classes, as well as the sound adaptation that occurs when affixes are created at affix boundary. For grammaticalization, drawing on data from Romance and Basque, it shows that the grammaticalization of auxiliaries can give rise to inflectional classes. Another major source is a cycle of two successive grammaticalizations of the same markers, which is shown to produce inflectional classes in Skou and Arapesh. Arapesh is peculiar in exemplifying a rare case of inflectional class through the linear ordering of affixes. Cases of sound change and grammaticalization produce systems of classes with very few classes. Sound change in particular partitions into two subclasses a preceding single class.

In contrast, when reanalysis gives rise to inflectional classes a larger number of classes are created at once, as is the case in Maori and Manam, as well as in Pama-Nyungan languages.

In addition to these three main sources, a number of new sources are identified and analysed, in particular the creation of inflectional classes through lexical strata in case of language contact, and morphological sources in the form of heteroclisis and deponency.

Finally, in two scenarios, it is shown that inflectional classes can arise from a previous system partitioning the lexicon into lexical classes. This is the case with systems of alienability distinctions in possessive paradigms, which often develop into systems of inflectional classes, although such systems have not been considered as inflectional classes previously. Alienability distinctions give rise to systems of inflectional classes presenting a large open class, opposed to a number of small, closed membership classes. Gender is another case where existing classes give rise to inflectional classes, if gender is marked directly on nouns together with another feature, most often number. This is shown to be the origins of systems of inflectional classes in Niger-Congo languages, as well as in Arapesh languages from New Guinea. In these cases, membership of the inflectional classes corresponds to the membership of each gender. Different sources for inflectional classes thus produce systems with different forms.

Acknowledgements

Writing a DPhil thesis is by many aspects a very solitary activity, but it also benefits from the help of a large number of people to whom I wish to express my gratitude. My first thanks go to my supervisors, Martin Maiden and Mary Dalrymple, who have been supportive and have contributed to making this work clearer way beyond their duties, in particular during difficult times when my health was a major trouble. Similar thanks are due to Lincoln College, and in particular Louise Durning, Senior Tutor, for her help and support.

Other scholars at Oxford have contributed to the theoretical discussion both about this thesis and other projects, most notably JC Smith, Sandra Paoli, Victor Acedo-Matellán, Louise Mycock, Wolfgang De Melo, Ash Asudeh, Rosalind Temple, Sam Wolfe, and Charlotte Hemmings. For years of ongoing discussion about the nature of autonomous morphology, and for her friendship, I am grateful to Louise Esher.

A number of researchers contributed to this work by providing data or discussion of it, in particular Emily Gasser and Laura Arnold. Yusuf Sawaki, Sonja Riesberg, Suriel Mofu and staff at CELD Universitas Negeri Papua in Manokwari were of great help during my fieldwork there, and in providing access a number of resources.

I am sorry to have too often bothered my friends with problems of inflectional morphology. Apologies and thanks go in particular to Pierre-Joan Bernard, Lara Ehrenhofer, Carryn Yong, Daniele Artoni, Skaiste Aleksandraviciute, Robin Meyer, Jean Quinsat, Marie Debax, Julie Teysseyre, Satrio Nindyo Istiko, Beatrice Rea, Gaspard Grandperret, Vincent Robert-Nicoud, Nathalie Raunet, Brendan Osswald, Élise Dupuy, Tom Juzek, Myriam Bras, Richard Parkinson, and Christopher Hollings. For his help during difficult times, I thank Benoît Levesque.

My grandparents Yvonne and Roger Aussenac, as well as my aunt and uncle Marie-Claire and Christian Carayon have always shown me love and support, and I could not thank them more.

Finally, this thesis would not have been possible without funding from the AHRC (doctoral award 13/44888), Lincoln College for the award of a Crewe graduate scholarship, and The Queen's College for employing me for two years as a Laming Fellow.

Table of contents

1. Introduction

1.1. Inflectional classes in diachrony

This thesis is concerned with the diachronic development of inflectional classes. It aims at understanding better the potential sources of those purely morphological phenomena that are inflectional classes, beyond the usual observation that they originate mostly in sound change.

Inflection is of great theoretical interest in that it appears unique to human language. To my knowledge, no formal or symbolic language makes use of a device similar to inflection (see Baerman 2015:2; Carstairs-McCarthy 2010). Inflectional classes, as lexical classes dividing the range of lexical items of a given language into sets of lexemes which inflect in a particular way, making use of different exponents depending on the classes, are all the more interesting because they are specific to natural languages, an unnecessary complexity.

To envisage the diachronic development of inflectional classes means examining the various diachronic sources that inflectional classes may have in language. The present study is thus structured according to such possible developments, envisaging each possible source in turn, and providing numerous examples of each development.

There has been a wealth of recent studies on inflectional morphology, including a number related to the issue of inflectional classes. These studies however tend to adopt a synchronic approach. They develop themes such as the correct way to formalize inflection and inflectional classes, or their description in some particular languages or language families. But very few studies are devoted to their diachronic development and to the origins of the appearance of inflectional classes, a notable exception being Dammel (2011).

Dammel's book is only concerned with Germanic languages, and is a study of the diachronic development of conjugation classes in these languages, which amounts to the study of inflectional classes restricted to a specific part of speech, in this case verbs. Dammel notes in her final synthesis that five types of developments can contribute to the origins of inflectional classes in Germanic

languages. Some of these are attested by historical records, such as the two types of development having to do with sound change, sound change proper, and reductive sound change due to the high frequency of some lexical items. Others can only be inferred from reconstruction, such as grammaticalization phenomena. To these two main mechanisms she adds reanalysis.

In this thesis I analyse these same three main mechanisms and extend the range of evidence to more diverse language families, adopting a crosslinguistic approach. Because the phenomena under investigation are not present in all languages, and because some of the sources are better exemplified by some languages than others, the range of languages is necessarily partial. My own limitations in the knowledge of the world's languages also skew the distribution of languages shown in this thesis. Some language families will feature more prominently because of my greater familiarity with them and with the main grammars and sources of data. These include the Romance languages, Basque, and Austronesian and non-Austronesian languages of New Guinea. I have also included data and analyses on languages from Africa, Australia, and North America, for some chapters. This is thus not a balanced sample of languages. For each possible development, I give a number of genetically diverse examples, in order to ensure that the possible source of inflectional classes under study is a common possible source.

The main idea underlying this thesis is that any type of development which is applied to a set of lexical items in the presence of inflection can give rise to inflectional classes. This is a very basic and broad idea, which is in no way a governing principle for their creation. Overall, it gives a rather heterogeneous survey, because the possible sources here identified are very diverse. In particular, I touch on the possibility, not envisaged by Dammel, that inflectional classes can originate in systems which already form a system of classes partitioning the lexicon: there are two possible cases developed here, inflectional classes originating in systems of alienability marked on possessive paradigms, and systems of gender when they end up being overtly marked on nouns.

Finally, it is impossible in such a study not to make some assumptions about the nature of human language, which will inform and constrain the approach taken here. In particular, I consider

that morphology is an autonomous component of grammar, on a par with phonology and syntax, but this by no means implies that there are not interactions with other components. This leads me to adopt a broad Word and Paradigm approach to morphology (Matthews 1991; Blevins 2016 among many others). In this approach, it is recognized that there are some purely morphological phenomena, including but not restricted to what some define as 'morphomes' in the form of specific distributions of root allomorphy, and also including inflectional classes.

As often as possible, I analyse surface forms rather than underlying forms. Phonologically underlying forms are always the result of a phonological analysis, and often presuppose that some alternations are accounted for by the phonology of the language. Depending on the depth of this analysis, using underlying forms might hide alternations which are important for the morphology. Some types of phonological analyses tend to account for these alternations only through phonological processes while some of those alternations might be recruted by the morphology to mark meaningful oppositions. It has been shown that surface forms, being what the hearer of a languages first receives, work better for the analysis of inflectional morphology (see Lass (1984) for arguments in favour of analysing surface forms; Blevins 2016). Finally, in a large number of cases, I depart appreciably from the analyses of the phenomena under scrutiny that are found in the descriptive grammars. This is most notable in the chapter about alienability distinctions, where I analyse such distinctions as inflectional classes, which is not their traditional analysis.

1.2. Structure of this work

The thesis is structured as follows. **Chapter 2** is concerned with defining what an inflectional class is, and analysing some of the issues with the extant definitions. After defining some key terms necessary for the understanding of a definition of inflectional classes, in particular the notions of feature, paradigm, lexeme and exponence, I discuss the main definitions in the literature, in particular the definition given within the framework of Canonical Typology by Corbett (2009). I

then turn to a number of issues related to the definition, mainly the question of the number of classes needed for it to be possible to talk of inflectional classes, and the number of members that each class should present to be considered a class and not simply an irregular lexeme. Stump (2016) makes the radically new proposal that inflectional classes should be understood as classes of stems, and not, as generally assumed in the literature, as classes of lexemes. After exploring his reasons for this proposal, I challenge it with both theoretical and empirical arguments. Inflectional classes should thus be understood as classes of lexemes. Finally, canonical inflectional classes should not be aligned on any extramorphological class, and in particular on word classes. I devote some time to showing that, for reasons having to do with the identity of content paradigm between inflectional classes, the sharing of inflectional classes between word classes can only arise in very exceptional circumstances. A number of cases of inflectional class sharing presented in the literature are shown to be erroneous, and some principles governing the possible sharing of inflectional classes between word classes are presented. These developments aim at giving a more precise understanding of the notion of inflectional class.

Chapter 3 is an overview. It briefly analyses all the possible sources for inflectional classes that I find in the literature or that I can derive from extant systems, some of which are developed further in the following chapters. Dammel (2011) identifies five types of possible origin for inflectional classes which in fact come down to three main mechanisms: sound change, grammaticalization, and reanalysis. Thus the chapter first develops these three main possible origins for inflectional classes, either illustrating them by simple made up examples for clarity of exposition, or with extant inflectional class systems. I then develop further possible origins for inflectional classes, in particular inflectional classes coming from a system of lexical strata in a situation of language contact or in creole languages; I also analyse the possibility for inflectional classes to originate in the morphological phenomena of heteroclisis (when a lexeme inflects according to two or more inflectional classes) and deponency (when the exponents for a given bundle of feature-value pairs are taken from the exponents for another bundle of feature value pairs elsewhere in the lexicon). I

then envisage the possibility for inflectional class to develop from systems of alienability distinctions in possessive paradigms, and from systems of agreement, in particular gender systems. The chapter is concluded by a reassessment of the No Blur Principle, which predicts that some types of inflectional classes should be impossible in natural languages. I show in particular that this principle does not hold for the Pama-Nyungan language Mpakwithi.

It has generally been assumed that the main origin for inflectional classes is sound change. When a language presents inflection for a given word class, sound change located in particular at the boundary between the stem and inflectional affixes can create alternations which are originally conditioned by phonology, but which can become morphologized and not reflect any particular synchronic phonological alternation. **Chapter 4** reassesses the different types of sound change that can give rise to inflectional classes. I examine instances of sound change which split existing inflectional classes, drawing on examples from Italo-Romance dialects, Romanian and Istro-Romanian. Some of those examples involve a process of metaphony which is originally a phonetic alternation in the stem vowel due to the assimilatory effect of a following high vowel, but which in many cases becomes morphologized as an inflectional marker. I then turn to instances of sound adaptation at an affix boundary which can create alternations between new classes of lexemes partially conditioned by the phonology of the stem. These are shown to be a source for inflectional classes in the South Halmahera-West New Guinea family of Austronesian languages.

Grammaticalization is another possible source of inflectional classes. In **Chapter 5** I show in particular that two main scenarios are possible for grammaticalization to give rise to inflectional classes, which have not been described as possible sources previously. The first has to do with the rise of auxiliation. An auxiliary construction will generally grammaticalize for only a subpart of the paradigm (a given screeve or screeves), and different auxiliaries can grammaticalize for different verbs. This is what happened in the Romance languages where two different auxiliaries present a lexically specified distribution between lexemes. Such a distribution of auxiliaries can be understood as an inflectional class. It is also possible that the auxiliary construction will not concern

all lexemes in the language: some verbs for example will inflect with an auxiliary bearing tense, mood and person information, while other verbs will inflect synthetically. This is the situation exhibited by Basque. Grammaticalization has been shown to operate for the adjunction of bound pronominals to verbs which become forms of agreement. In cases of successive cycles of grammaticalization of bound pronominals, different verbs may exhibit different markers, as it is the case in the two Papuan languages Skou and Arapesh.

Dammel's (2011) book on Germanic inflectional classes insists that reanalysis of the secretion type is a source of inflectional classes in these languages. **Chapter 6** analyses cases of reanalysis giving rise to inflectional classes in a broader range of languages, including the Austronesian languages Maori and Manam, some Pama-Nyungan languages, notably Yidiny and Nyawaygi, and in Romanian.

Chapters 7 and 8 examine two types of developments giving rise to inflectional classes from two other types of features, alienability distinctions in possessive paradigms, and gender in nominal paradigms. Such types of developments have not been treated before in the literature concerning the development of inflectional classes. Alienability distinctions are a partition of possessive paradigms between alienable items (those whose possession is felt to be more transient) and inalienable items (those whose possession is felt to be permanent). Typical inalienable items include kin terms, body parts and some cultural items such as names. **Chapter 7** first observes that a large number of languages deemed to present alienability distinctions in fact present a fixed list of lexical items whose semantics are not sufficient to classify them as inalienable or alienable. Some kin terms will belong to the alienable class, others to the inalienable. It is thus impossible to define the properties of the feature value 'inalienable' in purely semantic terms, but it has to be defined circularly as comprising those items which belong to the morphological class of inalienables. In addition, a large number of languages do not simply oppose two classes, but oppose more than two morphologically distinct classes. These classes are better modelled as inflectional classes which arise from the specific lexical semantics of some items, and the arrested grammaticalization of a new pattern

which does not affect very frequently possessed lexical items. That most alienability distinctions are in fact inflectional classes is further proven by the fact that a number of languages present heteroclite lexemes which inflect in parts of their paradigm according to the so-called inalienable pattern, and in other parts according to the alienable pattern. I provide crosslinguistic evidence for considering alienability distinctions as inflectional classes arising from specific types of grammaticalization from a number of language families, including Austronesian, Pama-Nyungan, Trans-New Guinea, Anêm (isolate), Pomoan, and Cochimi-Yuman.

Finally, **Chapter 8** develops the idea that some systems of inflectional classes marked on nouns may originate in a gender system. I first draw on the literature on the implication relations between gender and inflectional classes to compare and contrast the two types of systems. Then I develop the idea that a gender marker can regrammaticalize as a new marker on nouns, expressing at the same time values for the feature number or another feature (in all cases considered, it is number). I analyse how a system of gender may have given rise to a system of inflectional classes on nouns in two genetically and areally diverse language families, the Arapesh languages, and the Gur languages.

2. What are inflectional classes?

In this chapter, I provide a definition of inflectional classes, comparing the various definitions that have been given in the literature. The aim is to propose a clear and refined definition of the phenomenon, in order to ascertain that the systems of inflectional classes analysed in this thesis are truly comparable. I then analyse two related questions: how many lexemes are needed to make a class, and should inflection proper be modelled in a way similar to multiple class systems? This poses the question of the link between word classes and inflectional classes. The next section is devoted to a recent development in the literature about inflectional classes; Stump (2016) proposes that inflectional classes are not classes of lexemes as it is generally assumed, but classes of stems. His proposal is shown to be highly problematic for a theory of inflectional classes.

2.1. Prolegomena

Inflectional classes are classes structuring the inflectional realizations of a language in a particular way. Other types of classes also structure such realizations, for example the morphomic distributions which govern the distribution of stem alternants within a paradigm (Maiden 2005; Maiden et al. 2011; Cruschina, Maiden, & Smith 2013; Round 2015). In order to give a first definition of inflectional classes, some prerequisites and assumptions have to be made explicit, because a correct definition depends on a precise understanding of other notions, mainly those of INFLECTION, FEATURES, LEXEME, PARADIGM and EXPONENT.

For a language to have inflectional classes, it first needs to have inflection of some sort. The definition of inflection, often as opposed to derivation (and compounding, to some extent) has been the locus of much debate in the last two decades in morphological theory (Booij 1994; Booij 1996; Bauer 2004; Corbett 2009 among others). To give a first approximation, when defined one against the other, derivation opposes new lexical distinctions (new lexemes or oppositions between existing

lexemes), whereas inflection opposes word forms of a given lexeme according to some features required by the syntax or by a semantic feature value. As Baermann (2015:1) puts it, "Inflection is the expression of grammatical information through changes in word forms", whereas derivation is said to oppose new lexemic distinctions, that is new lexemes which can, in turn, be subject to inflection. A number of diagnostics have been suggested for distinguishing inflection from derivation, which vary in detail from author to author. Most have in common the five diagnostics suggested by Stump (1998:15-18 as summarized in Baerman 2015:3):

	Inflection	Derivation
Lexical meaning and/or part of speech	Same	Different
Obligatory	Yes	No
Productive	Yes	No
Semantically regular	Yes	No
Closure (recursivity)	Yes	No

Table 1. Diagnostics for inflection and derivation (Baerman 2015:3)

A general consensus seems to have been reached in understanding the distinction between inflection and derivation as a gradual distinction (Bauer 2004; Spencer 2013): there are clear-cut cases of both phenomena, but also a number of less categorical ones in between. Some categories for example can trigger derivation on some word classes but inflection in others: gender is such a feature, which in a lot of languages will be derivational on nouns, but inflectional on adjectives and on other agreement targets, as is the case in Greek.

Corbett (2009) provides a definition of inflection in terms of canonical typology, which is useful in that it defines inflection in its own terms, from a theoretical ideal, rather than opposing it to other phenomena. In canonical typology, "we extrapolate from what there is to what there might be, in order to define the theoretical space. Within that scheme of theoretical possibilities we can situate the real instances we find." (Corbett 2009). The canonical model first assumes that inflection is in fact a non-canonical deviation from the expected one-word one-meaning situation (completely analytic, no inflection, syntax only). For inflection, it also assumes that once features and their relevant values have been established, the paradigm should be exhaustive: there should be cells in the paradigm for all the possible combinations of feature value pairs for which the lexeme inflects

(the principle of *exhaustivity*, Spencer 2003). Thus if a language has a person feature, with three values (1, 2, 3) and a feature tense with two values (past, non-past), there should be six cells, each filled with a different form. The principle of exhaustivity in fact defines a content paradigm, the various combinations of values for which a given lexeme will inflect. For each combination there will be a word form. Canonical paradigms with respect to inflection are those paradigms which follow the principle in having all possible cells. Others are non-canonical in this respect.

The second principle is that of *consistency*, which follows the sub-criteria in Table 2:

	comparison across *cells* of a lexeme (level one comparison)	comparison across *lexemes* (level two comparison)
1. composition/structure (morphotactics)	same	same
2. lexical material (≈shape of stem)	same	different
3. inflectional material (≈shape of inflection)	different	same
outcome (≈shape of inflected word)	different	different

Table 2. Subcriteria for the principle of consistency (Corbett 2009:2)

These various criteria mean that formally, canonical inflection should exhibit a maximal differentiation in the exponents of each cell of the paradigm, while the lexical element should remain unmodified across all inflected word forms of a given lexeme. It also implies that different lexemes will have different lexical shapes, while still marking the different inflectional values in exactly the same form. If we abstract away from the effects of vowel harmony, a canonical situation would be that of nominal inflection in Turkish, as can be seen in Table 3:[1]

1 This is simplifying the situation in Turkish slightly, as some forms of the possessive paradigm are syncretic, which would infringe the principle of maximal differentiation of exponents. It still holds true for the basic, non-possessive nominal paradigm presented here.

	ADAM 'man'			TAVAN 'ceiling'	
	SG	PL		SG	PL
NOM	adam	adam-lar		tavan	tavan-lar
ACC	adam-ı	adam-lar-ı		tavan-ı	tavan-lar-ı
DAT	adam-a	adam-lar-a		tavan-a	tavan-lar-a
LOC	adam-da	adam-lar-da		tavan-da	tavan-lar-da
ABL	adam-dan	adam-lar-dan		tavan-dan	tavan-lar-dan
GEN	adam-ın	adam-lar-ın		tavan-ın	tavan-lar-ın

Table 3. Turkish paradigm for ADAM 'man' and TAVAN 'ceiling' (Stump 2016:31-33)

We can analyse these Turkish data according to the canonical situation described by Corbett (2009). As this paradigm inflects for two features, number and case, there are thus a maximum of two suffixes. Similarly, the principle of exhaustivity is satisfied: all 12 possible cells (two features with two and six values respectively) are filled with a different form. According to the first criterion of the principle of consistency, the morphotactics of the inflected word forms are identical for all forms both inside the paradigm of a lexeme and across paradigms. Every word form is composed of a stem followed by various suffixes, each realizing only one value of one feature (e.g. the value nominative of the feature case). As stated by the second criterion, the shape of the stem does not change for a given lexeme (*adam-*, *tavan-*), but is different for every lexeme: in theory, this states that cases of stem homophony in different lexemes are non-canonical. Following the third criterion, the shape of the inflectional material is different for every cell of a given paradigm (ACC.SG *adamı* vs ACC.PL *adamları* for example), but identical in a given cell for every paradigm (DAT.PL is marked by *-lar-a* for all lexemes). This means that exponence is consistent, and that there is a biunique relationship between the exponent and the grammatical feature it realizes: in effect, canonical inflection closely resembles the morphemic ideal (see Anderson 2015). In every case, the outcome is a different word form: in the canonical situation, no two word forms should be identical, even for accidental reasons. Thus a speaker encountering any of the word forms presented for Turkish can unambiguously deduce the feature value pairings expressed and isolate the lexical form to be retrieved and interpreted.

An Austronesian example of a system of inflection quite close to the canonical ideal can be

found in Mangap-Mbula (North New Guinea, Oceanic; Bugenhagen 1995:117)[2]. Mangap-Mbula only inflects verbs for the features NUMBER with values singular and plural, and PERSON with values 1, 2, 3, with an additional distinction in terms of CLUSIVITY for first person plural forms. What makes it slightly non-canonical is the fact that this paradigm does not respect the principle of exhaustivity: CLUSIVITY is here a restricted feature only applying to first person plural[3]. The fact that some markers conflate the realization of more than one feature value is not a problem in terms of canonical inflection, as each set of values is still realized by a unique exponent, although these exponents are non-canonical in showing cumulative exponence (see below). Table 4 gives the prefix markers used in inflection and an example of an inflected verb:

FEATURES/VALUES	Prefix	*kam* 'do/get'
1SG	aŋ-	aŋkam
2SG	Ø-[4]	kam
3SG	i-	ikam
1PL.INCL	t-	takam
1PL.EXCL	am-	amkam
2PL	k-	kakam
3PL	ti-	tikam

Table 4. The verb paradigm of Mangap-Mbula (Bugenhagen 1995:117)

A LEXEME is an abstract construct which is often equated informally with an entry in the dictionary. It is in fact one such entry in the database formed by the mental lexicon, and would be more properly defined as a network of qualitatively distinct information. I will follow here the broad definition given by Spencer (2013) which defines the lexeme as a network of data of four different types. Minimally, a lexeme comprises data on the phonological form of the word, its word class, and thus its syntax, as well as information about its morphology and semantics. In addition,

2 I simplify somewhat the system for reasons of exposition. The language has in fact both inflecting verbs (which all exhibit the same, differentiated markers) and non-inflecting verbs, which would probably fall under a system of multiple classes. Note that underlying clusters in this language are broken up in surface realizations by addition of an epenthetic vowel copied on that of the stem.

3 Clusivity is in that respect a non-canonical feature, in that it cannot be present outside cells which realize first person and non-singular number, by definition. If the non-canonicity is a fact about the feature (limited in essence to 1PL in Mangap-Mbula), then the principle of exhaustivity could be said to hold.

4 One dialect has a further distinction here with two different markers depending on the phonological context. For all other dialects, the stem remains unmarked.

there will be a unique piece of data identifying the lexeme uniquely as separate from other lexemes, much in the guise of a primary key in a database. Very often this key is represented as the lemma, which conventionally is represented as the basic word form written in capitals. Thus the lexicon entry for the lexeme PLAY will involve four types of information. First, there will be some phonological information, which will define how the word should be pronounced. This is basically a phonological representation, but it can become more complex when a word has many different stems, or in case of suppletion. In the case of PLAY this could be represented as /plɛj/. Then, there will be information about the meaning of the lexeme, in this case a sort of equivalent of the definition in the dictionary. A syntactic category will be given to the lexeme, such as noun, verb, with specifications about its syntax, for example whether a verb is intransitive, transitive, ambitransitive or ditransitive. Finally, there will be information about its morphology, which amounts to the inflectional paradigm it inflects for. In the case of PLAY as a verb, this will be the basic paradigm of verbs in English. Changing one of these four levels of data will involve creating a new lexeme with a specific type of relationship to the previous lexeme. Thus changing only the phonological representation will involve a synonymic relationship: two lexemes which have a different phonological representation, but the same word category, the same morphology, and crucially the same meaning. Similarly, two lexemes may have the same phonological representation but different semantics. In this case one talks of homonymy. Or two lexemes may have similar meanings and phonological representations, but a different word class (and thus a different morphology), in which case one talks of conversion, as in the case of the conversion of a noun into a verb in English, for example 'Google' and 'to Google'. These types of relationships may be understood as non-canonical relationships between lexemes which canonically should differ on every level of the representation.

FEATURES can be defined informally as those grammatical meanings and indices realized either through inflection or by a distinct word in different languages. In English for instance, there is a

distinction encoded through inflection between *cat* and *cats*, which is similar to the distinction between *ox* and *oxen* in that one form refers to a single entity in the world, whereas the other refers to multiple such entities, although the formal means of expression are different. In this case, it can be said that English makes a distinction for the feature NUMBER between two distinct values, SINGULAR and PLURAL. Establishing an inventory of such features and their relevant values is a prerequisite for the analysis of any inflectional system. Corbett (2012, 2015) establishes an inventory of features, and creates a typology of features which I follow here. He distinguishes in particular features which are specific to one component of the grammar, such as phonological features, and do not play a role in the other components, and interface features which play a role at the interface of two components of the grammar. Interface features can be for example morphosyntactic features, which mediate between syntax and morphology (Corbett 2015:35–36; Corbett 2012). Concerning inflectional morphology, four types of features are distinguished: morphosemantic, morphosyntactic, morphophonological, and purely morphological (or morphomic) features.

MORPHOSEMANTIC FEATURES are features which bear no role in the syntax, but relate semantic values to a specific word class for which they are relevant. NUMBER is such a feature when applied to nouns: the speaker makes a meaningful choice in choosing to refer to one or more than one entity; thus expressing a noun in the singular or in the plural is a meaningful choice by the speaker, a matter of semantics, and does not a priori bear any direct consequence for the syntax. Similarly, TENSE applied to a verb categorizes the event referred to by the verb according to its location in time, and VERBAL NUMBER categorizes the verb for the number of events referred to. This means that such features are generally linked to a specific word class, in that they realize features that are expected for this type of word class. On the other hand, MORPHOSYNTACTIC FEATURES do have consequences for the syntax: the fact that English verbs encode a feature NUMBER has nothing to do with the inherent properties of the event referred to by that verb, but with the number value of its syntactic subject, through a rule of agreement (Corbett 2015:36–37). It is thus an agreement feature.

As noted by Corbett, morphosyntactic features should be distinguished from morphosemantic features for conceptual as much as typological reasons, as they "are more limited in number, and they tend to correspond more closely across languages" (Corbett 2015:37). The distinction made by Corbett between these two types of features is rather reminiscent of the distinction between inherent inflection and contextual inflection (Booij 1994; Booij 1996). As stated by Booij: "Inherent inflection is the kind of inflection that is not required by the syntactic context, although it may have syntactic relevance. Examples are the category number for nouns, comparative and superlative degree of the adjective, and tense and aspect for verbs. (...) Contextual inflection, on the other hand, is that kind of inflection that is dictated by syntax, such as person and number markers on verbs that agree with subjects and/or objects, agreement markers for adjectives, and structural case markers on nouns." (Booij 1996:2)

This amounts to saying that inherent inflection reflects the realization of morphosemantic features, whereas contextual inflection is the formal realization of morphosyntactic features. But distinguishing the two types at the level of features has enormous advantages in terms of clarity and formalization: in cases where multiple feature values are realized in a single cell through cumulative exponence (many feature values realized as one form, see below), it is unclear how a distinction of two types of inflection could be made at any other level than features. In English for example, a suffix *–s* realizes on verbs the feature value pairs {<NUMBER:SG>, <PERSON:3>, <TENSE:PRES>}, including features which would characterize this type of inflection as both inherent and contextual. At the feature level, though, only the TENSE feature is morphosemantic (inherent), the other features being morphosyntactic features expressing agreement (contextual). The distinction is also important for theories of agreement: controllers of agreement tend to realize morphosemantic features which are matched through agreement with corresponding morphosyntactic features on the target.

Just as morphosyntactic features are interface features between morphology and syntax, MORPHOPHONOLOGICAL FEATURES are interface features between morphology and phonology. They

involve phonological alternations which do not follow the general rules of phonological alternations in the language, but which are to some extent both morphologically constrained and relevant to morphology, as they may encode specific patterns (in a similar manner as stem selection patterns), and exhibit implicational relationships between patterns. Some theories would treat such alternations directly in the phonology, because they essentially involve phonological alternations of the type usually modelled by phonology, although these may be lexically specified, or linked to a specific morphological environment or class. A theory positing cophonologies (positing that a given language has a number of different phonologies whose expression is constrained by external factors such as word class, inflectional domain, etc.) would for example treat these alternations as a phonological subsystem limited to a given word class (see, e.g., Inkelas 2014). In most instances, such restricted alternations are in fact remnants of former phonological alternations which have ceased to be productive elsewhere in the language, and can help reconstruct previous stages of it (Corbett 2015:44). The important point here is that such morphophonological features are relevant to the morphological structure of a language. In particular, if such alternations can become fossilized inside words while the general rules that gave rise to them are no longer applicable in the language in synchrony, it calls into question the analysis of alternations that follow general rules of phonology as not playing a role in morphology as well (see Maiden 1991).

Finally, MORPHOLOGICAL FEATURES in fact largely correspond to the notion of MORPHOME first proposed by Aronoff (1994), and further developed since (Maiden 2005; Maiden et al. 2011; Cruschina, Maiden & Smith 2013; O'Neill 2011; O'Neill 2014; Round 2013; Round 2015 among others), and are sometimes referred to as morphomic features. These are features only relevant to the (inflectional) morphology, which relate specific items in the lexicon with a variety of patterns. Aronoff first proposed that inflectional classes and stem selection patterns should be considered as morphomic (Aronoff 1994). Other types of morphological features have since then been proposed, which Corbett analyses as less prevalent, some of which may be restricted to a limited set of languages and potentially analysable in terms of a combination of morphological and

morphophonological features. Because such features are the main object of inquiry in this thesis, their definition will be developed in the next section, centring on inflectional classes.

EXPONENCE is the formal realization of a set of feature values. In this sense, an exponent is some phonological material that formally expresses or signals the features in inflection. Two main typologies of exponents have been proposed, based on a different view of the relationship of exponents with features on one side (e.g. Matthews 1972), and with what might be considered the base form of a lexeme on the other (Trommer 2012). One important assumption that is often made is that no type of exponence is specific to inflectional morphology (see Trommer & Zimmermann 2015:47). The relationship between exponents and features can be devised in a canonical typology framework: the canonical exponent corresponds to the canonical morpheme, understood as some piece of phonological material (or process) systematically correlated with the realization of one and only one feature value. This is the situation found in the nominal paradigm of Turkish discussed above for canonical inflection: every exponent is correlated with only one feature value, such as the exponent *-lar* realizing a feature value NUMBER:PL, or the exponent *-a* realizing a feature value CASE:DAT. Non-canonical types of exponence can be defined from that ideal of a 1:1 mapping of form and function, as can be seen in the following table:

Exponent	Feature value pair	Type of exponence
1	1	Canonical exponence
many	1	Extended exponence (Matthews 1972:50), or distributed exponence
1	many	Cumulative exponence (Matthews 1972:72-75)
many	many	Overlapping exponence[5]
0	1/many	Zero exponence
1/many	0	Empty morphs

Table 5. Exponence relations

Another important distinction to be made about exponence concerns the question of linear ordering. Even in additive exponence (exponence which takes the form of affixes), the position of the exponent in the linear ordering of the word form matters: some features will be realized preferably by prefixes, suffixes, infixes, or circumfixes, and similarly in cases of non-additive exponence, linear ordering can matter, for instance in the case of exponence through initial consonant mutation (see below the case of Mali).

The second typology of exponence starts with the assumption that exponence is best formalized as operations on a base form (Trommer & Zimmermann 2015), which 'express' some inflectional category. Trommer and Zimmermann thus distinguish additive exponence, characterized by the realization of inflectional feature values by the addition of some phonological material to the base form of a lexeme; transformational exponence, whereby these inflectional feature values are realized by the transformation of the phonological material of the base form of a lexeme; and templatic exponence, when the realization takes the form of an adjustment of the phonological

5 This is in theory a possibility: multiple feature value pairs are realized through multiple exponents without any possibility of assigning one exponent to any of the feature value pairs. Such a possibility is given by Ancient Greek examples cited by Peter Matthews, where multiple feature values are realized by exponents which also realize other feature values:

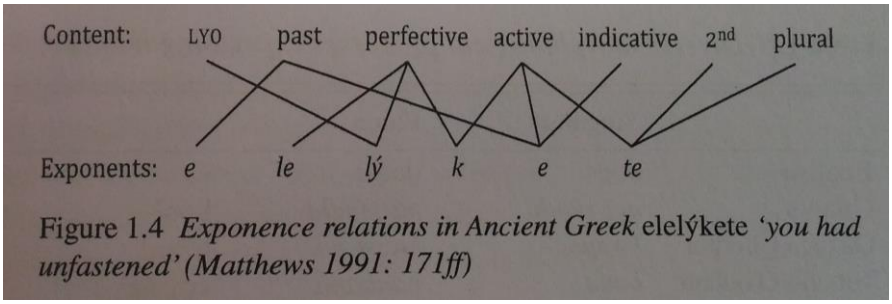


Figure 1.4 *Exponence relations in Ancient Greek* elelýkete *'you had unfastened' (Matthews 1991: 171ff)*

That distinction is never made explicitly in the literature, but I suggest using the term 'overlapping exponence', used by Stump as a cover term for all cases of cumulative and extended exponence, for this specific case (see Stump 2016:21).

material of the base to a fixed shape (Trommer & Zimmermann 2015). Both typologies have their advantages, in that they do not take the question of exponence from the same angle: in one, the question is looked at from a morphosyntactic point of view, considering exponence in its relationship with the feature values to be realized; in the other, a phonological point of view is taken, considering the phonological shape and rules needed to give form to the realized features.

A PARADIGM may be defined as the complete inventory of forms taken by a single lexeme for its relevant feature value bundles.[6] Some theories view paradigms essentially as an epiphenomenon, having at best pedagogical value: this is mostly the case with theories making use of morphemes for the encoding of inflectional features. In the Word and Paradigm family of theories, whose perspective is taken in this thesis, paradigms are seen as central in the organization of the inflectional system of a language, mostly because of "the general insight that one inflection tends to predict another" (Matthews 1991:197). This means that paradigms are characterized by specific organization rules, making in particular use of principal parts, which are those cells that can help predict the form of other cells in the paradigm of a given lexeme. Word-based morphology approaches (Blevins 2005; Blevins 2006) tend to view paradigms as "cell-form pairs that are related not by shared bases or derivational histories but networks of interdependencies" (Blevins 2015:89). Other models tend to treat paradigms as primarily n-dimensional matrices where each cell is treated as an abstract feature value pair bundle, and where each dimension of the matrix is defined by one of the inflectional features of the system. This is the case with Paradigm Function Morphology (PFM, Stump 2001), and more generally of most realizational theories of inflectional morphology.

[6] Originally and etymologically, a paradigm is a model, and the term should thus be reserved for those exemplars which serve as model for a larger number of lexemes. In that sense, it could not be used for irregular lexemes. But in recent works in morphology, the term seems to be used more largely for any model of inflection, whatever the number of lexemes which may follow this model, and even for irregular lexemes.

2.2. Defining inflectional classes

Aronoff defines the notion of INFLECTIONAL CLASS as "a set of lexemes whose members each select the same set of inflectional realizations" (Aronoff 1994:64). In a very simple case, which corresponds to the canonical ideal, this could give the following system, for lexemes only inflecting for a feature number, with two values SG and PL:

	CLASS 1	CLASS 2
SG	aba-ri	beda-mi
PL	aba-ra	beda-ma

2.2.1 Some examples

We can start with two relatively simple (though quite different) examples of inflectional classes: Biak, an Austronesian language, and Mali, a Baining language, both spoken on islands close to New Guinea. Though these two languages have a limited inventory of features in their verbal system, it is still quite complex to decide a priori on the number of classes that should be recognized. They also show very different patterns of exponence: in Biak, most features are realized with affixes (prefixes mostly, and some infixes), whereas Mali mostly makes use of initial consonant mutations on the verb stem.

Biak is an Austronesian language of the South Halmahera-West New Guinea family, spoken on a wide range of islands around the Bird's Head of New Guinea, including Biak, Numfoor, Waigeo and parts of the mainland (Steinhauer 2005; van den Heuvel 2006; Mofu 2008). Table 6 presents the relevant paradigms for the three classes of verbs, in the Biak Betew dialect as spoken on Mutus island (own field notes):

	PRO	Class 1	Class 2	Class 3
		'wake up'	'come'	'sleep'
ROOT FORM		KWOEK	RAMA	ENEF
1SG	aya	yakwoek	yarama	yenef
2SG	au	wakwoek	rwama	wenef
3SG	i	ikwoek	ryama	denef
1DU.INCL	ku	kukwoek	kurama	kuyenef
1DU.EXCL	nu	nukwoek	nurama	nuyenef
2DU	mu	mukwoek	murama	muyenef
3DU	tu	tukwoek	turama	tuyenef
3PAUC	tko	tkokwoek	tkorama	tkenef
1PL.INCL	ko	kokwoek	korama	kenef
1PL.EXCL	nko	nkokwoek	nkorama	nkenef
2PL	mko	mkokwoek	mkorama	mkenef
3PL.ANIM	si	sikwoek	trama	tenef
3PL.INAN	na	nakwoek	ndrama	nenef

Table 6. Biak Betew pronouns and verbs

Biak only marks agreement with the subject on the verb, that is, only the morphosyntactic features person with values 1, 2, and 3; clusivity with values inclusive and exclusive; number with four values, singular, dual, paucal and plural; and finally gender, with the values animate and inanimate. There is a restriction in that certain combinations are limited: paucal value for number can only combine with third person, and the gender feature only combines with third person plural. This means that such a paradigm does not satisfy the criterion of *exhaustivity* posited earlier for canonical inflection: not all the possible combinations of feature values are attested in the language. Class assignment is partially conditioned by phonology in Biak: all CC-initial verbs belong to class 1, all V-initial verbs belong to class 3. Both of these classes realize inflectional features by means of prefixes, and, to some extent, the variation between the two classes could be explained in terms of morphophonology, except for the form realizing 3PL.ANIM where class 1 *si-* corresponds to class 3 *t-*[7]. First and second class only differ for a few cells (2/3SG, 3PL), and both are formed of C-initial lexemes, for which class membership has to be lexically specified. For these lexemes, only four cells can function as principal parts, those cells that can predict the inflectional class membership of

[7] This is initially a phonological alternation in this dialect of Biak, where /s/ > [t] in a number of environments, but retains [s] in front of high vowels and glides. It has since morphologized because of the loss of /j/, which did not trigger an allophone [t].

a lexeme. Class 2 verbs are characterized by a different linear ordering of the feature value realization for 2/3SG, where inflectional values are realized with an infix. The system is thus relatively simple, but still highly non-canonical in its realization: classes are far from being maximally differentiated, the assignment to an inflectional class is not purely lexical but partially conditioned by formal specifications (phonology), and the paradigm is not exhaustive.

Mali (Stebbins 2011) is a Baining language spoken in the East New Britain province of Papua New Guinea. It has the particularity of only encoding a TENSE feature by means of inflection on the verb.[8] The formal realization of the tense distinctions is not affixal, but involves a quite complex system of consonant mutations, as can be seen for the four classes of verbs presented by Stebbins (2011:51-53):

Present	Past	Future	Gloss
dēn	muēn	thēn	arrive
tes	mes	thes	call, read

Table 7. Class A verbs in Mali (Stebbins 2011)

Present	Non-present	Gloss
bang	vang	run
prap	vrap	fly
valing	waling	turn
dong	thong	turn
thoret	roret	crawl

Table 8. Class B verbs in Mali (Stebbins 2011)

Class A and B both involve complex consonant mutations for the realization of tense values, but they differ essentially in the values available, according to Stebbins: class A verbs distinguish three values, whereas class B verbs mark an opposition in terms of present vs non-present, a type of tense system that is extremely rare crosslinguistically. Class C verbs mark a similar opposition, but through affixation either in the present or in the non-present. If one of the forms were to be

8 Subject and Object reference are encoded with proclitics and enclitics respectively, according to Stebbins (2011:44-46). It is unclear why she treats these as clitics, as they always appear adjacent to the verb, take part in the interpretation of tense, show allomorphy linked to the inflectional class of the stem and sometimes extremely idiosyncratic phonetic reduction with stems such as *thet* 'go'. I chose here to treat only the verb system as Stebbins (2011:51-53) presents it.

considered as more 'basic', the choice would either have to be dynamic and dependent on the class (maybe for reasons of local markedness), or the system would exhibit both cases of exponence through affixation and subtraction:

Present	Non-present	gloss
tlu	lu	see
tnok	nok	cry
ngim	ingim	search
ngip	ingip	die

Table 9. Class C verbs in Mali (Stebbins 2011)

CLASS D verbs are verbs which do not show inflection for tense, such as *su* 'try'.

Another analysis is possible here, which does not exactly change the number of classes suggested by Stebbins, but which is more consistent with the theory of features and values developed by Corbett (2012). Because some verbs make a clear distinction between three tense values (present, past, and future), there is no principled reason to consider that other verbs do not make such a distinction, with use of extensive and systematic syncretism between past and future forms for types B and C, and between all cells for type D. This has the advantage of presenting a unified system of feature value pairings for all verbs. Deciding whether a given language presents two concurrent features of the same type or only one with a wider range of values is not trivial, but even in well-known cases such as split ergativity unitary analyses have been proposed that deal with the compleity of the systems without presenting two concurrent systems (or in the case of Mali, three) (Goddard 1982 on case systems; Round & Corbett 2017 for a similar analysis on TAM systems). After all, this is a similar case to the person and number distinction of English present tense: one could analyse a vast majority of verbs as having only two cells, one expressing third person singular, the other non-third person singular but evidence from the conjugation of the verb *be* and evidence from the pronominal system leads to the recognition of three values for person and two for number in all verbs, with a systematic pattern of syncretism for all but one. A similar case occurs in Romanian where there is complete syncretism between the present indicative and the subjunctive in all non-third person forms, except for the verb 'be' which distinguishes them in all

persons. On the basis of the sole verb 'be' a distinction between present and subjunctive is still assumed for Romanian. If one adopts such an analysis, it becomes clear that patterns of syncretism define three major classes, with the second class (syncretic between past and future) presenting a number of subclasses:

Present	Past	Future	Gloss
dēn	muēn	thēn	arrive
tes	mes	thes	call, read

Table 10. Class 1 verbs in Mali: no syncretism

Present	Past	Future	Gloss
bang	vang	vang	run
prap	vrap	vrap	fly
valing	waling	waling	turn
dong	thong	thong	turn
thoret	roret	roret	crawl
tlu	lu	lu	see
tnok	nok	nok	cry
ngim	ingim	ingim	search
ngip	ingip	ingip	die

Table 11. Class 2 verbs in Mali: systematic past-future syncretism

Present	Past	Future	Gloss
su	su	su	try

Table 12. Class 3 verbs in Mali: syncretism present-past-future

Biak and Mali show that the definition of inflectional classes rests on a correct definition of features and their values, which amounts to saying that it is dependent on a clear definition of the content paradigm of a group of lexemes. Under Stebbins' analysis, Mali does not present inflectional classes, because various groups of verbs present very different content paradigms. Under a different analysis, Mali presents inflectional classes, because all verbs share a content paradigm, and only differ in their form paradigm, because of extensive syncretism defining different inflectional classes.

2.2.2 Canonical inflectional classes

The idea of a canonical typology proposes that one should first define a given feature in abstract terms, establish an ideal canon of what that feature is, so that extant systems in real natural

languages can then be compared to the canon with regard to their variation to the ideal. It is rooted in the first attempts at defining a canonical typology of vowels by Daniel Jones: cardinal vowels are not expected to be found in real languages, or at least not all of them; but the idea of positing cardinal vowels as a canonical ideal enables the researcher to then place real languages in a typology with respect to the deviations from the ideal which are instantiated in real languages. As such, one does not expect to find the canonical ideal: it is thus quite different from the idea of a prototypical system. The notion of canonical inflectional class has been developed by Corbett (2009). Canonical inflectional classes are defined according to nine criteria which are grouped under two general principles, a principle of *distinctiveness* stating that "canonical inflectional classes are fully comparable and are distinguished as clearly as is possible", and a principle of *independence* which states that "the distribution of lexical items over canonical inflectional classes is synchronically unmotivated" (Corbett 2009:2–5). I return to the second principle below, but it is already important to note that in the canonical situation nothing prevents a system from being motivated in diachrony at some point of its development. I will return to that point in the course of the thesis and in particular examine whether these 'motivations' are at the origins of actual inflectional class systems: actual inflectional systems generally show a certain degree of motivation on one of the dimensions distinguished by Corbett, which give an indication about their origin. The principle of distinctiveness is characterized by the following criteria:

Criterion 1	In a canonical situation, forms differ as consistently as possible across inflectional classes, cell by cell
Criterion 2	Canonical inflectional classes realize the same morphosyntactic or morphosemantic distinctions (they are of the same structure)
Criterion 3	Within a canonical inflectional class each member behaves identically
Criterion 4	Within a canonical inflectional class each paradigm cell is of equal status

Table 13. Criteria for the principle of distinctiveness (Corbett 2009:4-5)

The criteria grouped under that first principle of distinctiveness are retained as being the most important for defining canonical inflectional classes (as in Stump 2015 for example).

The principle of independence is characterized by the following five criteria:

Criterion 5	The larger the number of members of an inflectional class (up to an equal 'share' of the available items) the more canonical that class
Criterion 6	In the canonical situation, the distribution of lexical items over inflectional classes is not phonologically motivated
Criterion 7	In the canonical situation, the distribution of lexical items over inflectional classes is not syntactically motivated
Criterion 8	In the canonical situation, the distribution of lexical items over inflectional classes is not motivated by Part of Speech
Criterion 9	In the canonical situation, the distribution of lexical items over inflectional classes is not motivated by pragmatics (including information structure)

Table 14. Criteria for the principle of independence (Corbett 2009:6-7)

This second principle is in fact divided into two types of criteria. Criterion 5 concerns inflectional class membership from a quantitative point of view, looking at the number of lexemes belonging to a given class. It concerns more the degree of canonicity of a specific class within an inflectional class system than the canonicity of a system as a whole, except when considering that a canonical system would have an equal number of members in each class. Any system with unequal classes would be non-canonical, whether it is because a given class has more members than canonically posited, or because it has fewer. The remaining criteria are concerned with inflectional class assignment: in a canonical ideal, the membership of a lexeme in a given inflectional class should be purely lexically specified, and no external factor should determine its membership to a specific class. Interestingly, Corbett's criteria 6 to 9 all define a similar situation: canonical inflectional classes should not appear as 'motivated' or as 'depending' on any other level of the grammar. This makes sense in the canonical ideal, as inflectional classes are defined as purely morphological structures, but there is little chance not to see any interactions with other components of the grammar, for the historical reasons I develop in the thesis: inflectional classes come from distinctions having to do with phonology, gender, semantics in diachrony, and in most cases the present-day situation shows remnants of this past systematic link with other levels of the grammar. A second interesting fact is that these four criteria define what one could subsume under the notion of covert and overt inflectional classes, borrowing the distinction from recent work on stem distributions (Smith 2013; Esher 2014): canonical inflectional classes should be overt in that it should be clear from any form of a given lexeme to what class it belongs and that the choice of a

given exponent should not appear to be conditioned by extramorphological factors, whereas a number of real instances can appear to be at least partially covert, which means that they can seem to be partially conditioned by extramorphological values. Class assignment in Biak is non-canonical in that respect, as the phonological (segmental) shape of a root largely constrains the assignment to a class: a root starting with a consonant cluster is assigned to class 1, a root starting with a vowel to class 3, and the remaining lexemes are distributed between classes 1 and 2. But even in cases where class assignment would seem to be purely conditioned by phonology, the alternation between sets of markers is morphological in essence: this is a case that I call covert.

Criterion 8 is surprising: one would expect inflectional classes to be linked strongly to a specific part of speech (or word class), at least if one considers that some features of inflection are more closely linked to some word classes than to others. This is in particular the case when some word classes inflect for morphosemantic features, such as number for nouns, verbal number on verbs (referring to the number of the event and not to the number of participants) or tense and aspect for verbs. Because canonical inflectional classes should be of the same structure (i.e. inflect for the same feature value pairs, from Criterion 2), this criterion would predict that systems where morphosemantic features are realized through inflection could not give rise to canonical inflectional classes[9], though Corbett himself states that because one finds some instances of inflectional classes that crosscut word classes in Russian,[10] and although this criterion will by and large not be met in most instances of inflectional classes, it should still be maintained for canonical inflectional classes (Corbett 2009:7). Some models assume a clear link between inflection and word classes (e.g. Network Morphology: Brown & Hippisley 2012; Stump 2015).

Stump (2015:114-115) for example defines a canonical situation for inflection as presenting an isomorphy between a word class, a content paradigm associated with all the members of that word

9 The reason is that morphosemantic features are linked to word class, and so sharing inflectional classes could not happen between word classes if morphosemantic features were expressed through inflection. I return to this problem below when examining under what conditions inflectional classes can be shared across word classes.

10 Another frequently cited example is Latin nouns and adjectives, which share two of the five inflectional classes present for nouns, and the a-stem inflectional class of nouns in Sanskrit, to which ordinary adjectives are also ascribed. I return to this question in the section 2.4. below.

class, and a form paradigm associated again with all members of that class. Inflectional classes, in his definition, only arise when some of the lexemes, although associated with the same content paradigm (set of feature value pairs for which a given lexeme inflects), make use of a different set of realizations (form paradigm), at least partially. In fact, that definition leads one to recognise that there is an inflectional class as soon as one form deviates for one lexeme. Interestingly enough, Stump only defines inflectional classes as a property of the form paradigm for a given word class. When two lexemes from the same word class present a different content paradigm, this is analysed as instances of defectiveness or overdifferentiation (see Anderson 2015 on the distinction between these two notions). Because the association of a given content paradigm with an entire word class is not systematic, it is not clear that a difference in content paradigm should not be counted as an instance of inflectional class: it will, after all, be reflected in the set of exponents, as there will still be a difference in the forms used by two different classes.

2.2.3. How many classes are needed to have inflectional classes?

The architecture needed for formalising one class is no different from that needed for formalising many. One could thus consider that inflection proper is just another instance of inflectional class, but one that is linked to the entire word class in an implicational relationship.

A number of works on inflectional classes do not precisely define the notion, although it is central to their analysis of inflectional morphology (Wurzel 1987 among others, but see Wurzel 1989 for a complete definition); other works do define inflectional classes, but their definition is not always consistent with their use of the term (Wurzel 1989; Aronoff 1994). Authors seem to agree on the fact that to speak of inflectional classes, one needs at least two (Corbett 2009 is the clearest example; also see Wurzel 1987, Aronoff 1994, Stump 2001, 2015, Brown & Hippisley 2012).

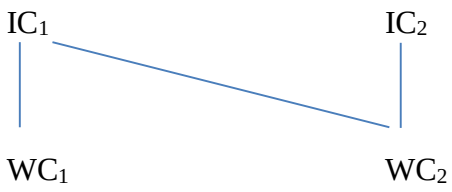
Wurzel (1989:51-59) spends some time defining three basic concepts of inflectional systems, markers, morphological categories (i.e. features and their values) and inflectional classes. For him, it is relatively straightforward to define inflectional classes, in that they are purely morphological classes where "no reference need be made to extramorphological criteria of 'content'" (Wurzel

1989:56; but see below the contradictory definition of Stump 2015). For Wurzel, the following criteria apply:

- "A language L_i has an INFLECTIONAL CLASS IC_j if, for a word group WG_j, every derived morphological category C_k (every derived categorial bundle C_{k1}... C_{km}) is symbolized by the same form, and all the derived inflectional forms are formally distinct from all the derived inflectional forms of all other word classes. There is a case of distinctiveness if and only if:
    - (i) at least one category C_l (categorial bundle C_{l1}.... C_{lm}) is symbolized so that it is phonologically different from the symbolizations of all other groups of words in C_{l1} (C_{l1}... C_{lm});
    - (ii) the symbolization of a category C_l (categorial bundle C_{l1}... C_{lm}) is phonologically identical with the word group WG_p and distinct from the word group WG_q as well as category C_r (categorial bundle C_{r1}.... C_{rm}) which is phonologically identical with the word group WG_q and distinct from the word group WG_p." (Wurzel 1989:56-57)

Wurzel here takes similar assumptions as canonical typology. Inflectional classes rely on a similar content paradigm, but they differ in the marking of the feature-value pairs contained in that paradigm. The important thing here is that Wurzel assumes that to have an inflectional class, one needs two at the very least, but he does not assume that the inflectional class should be related to word class: as long as the content paradigm is identical, there is no objection to words belonging to different word classes taking a similar inflectional class. In his own terms, "the words of such groups often fall into one word class, but they need not do so" (Wurzel 1989:57; see below for examples in Latin). The notion of derivation is of importance here in that it assumes, contrary to most assumptions in Word and Paradigm models, that there is a basic form of a given lexeme (an 'unmarked' form), whose marking can vary without the inflectional class being changed. Thus Wurzel classifies as belonging to one and the same class Latin nouns such as *equ-us* and *vir*, although the realization of the nominative singular is quite different in both, under the assumption that nominative singular is the 'basic' or 'unmarked' form in the paradigm, and that all other derived

forms make use of an identical set of affixes for both nouns (Wurzel 1989:58). Similarly, features are determined by the fact that they appear in at least one inflectional class, such as case for nominals in English on the basis that pronouns also mark it. I am doubtful that this should be done, on the basis that these are two very different word classes, and that the determination of word classes and inflectional classes should probably be kept separate. One (possibly unwanted for Wurzel) result follows from his definition: one can imagine a situation where, for a given word class, there is only one inflectional class (recognized as such), so long as it is shared with another word class presenting two inflectional classes for an identical set of feature-values. This would give the following system:



Thus, albeit in a very specific case, Wurzel's definition of inflectional class allows for a given word class to still present an inflectional class when there is only one class. But in other works, he still seems to be working under the assumption that at least two inflectional classes are needed (Wurzel 1987 analysing Turkish nominals).

Aronoff's (1994) definition allows more generally for a single class of inflectional markers to be considered as an instance of inflectional class. For him, "An *inflectional class* is a set of lexemes whose members each select the same set of inflectional realizations" (Aronoff 1994:64). Thus any set of lexemes inflecting in the same way could be said to be an inflectional class, including the case where there is only one such set for a given word class (or even for multiple word classes marking the same feature-values). If all nouns in a given language inflect with the same set of exponents, it is simply the case that out of the set of words, only those which are members of the set nouns can take these exponents. Again, no systematic relationship is being made by such a definition between word class and inflectional class: it is perfectly possible for the two levels not to coincide. When they do coincide, it is a case of (bi)implicational relationship between word class

and inflectional class ($LEX_N \leftrightarrow IC_N$). In practice, Aronoff considers that at least two such sets of lexemes are needed for a given word class to have inflectional classes; he is thus making both a link between word class and inflectional class, and with the fact of presenting two as necessary (Aronoff 1994:64). But in a note, he clearly states that, strictly speaking, one class is still an instance of inflectional class:

> "Strictly speaking, a language whose major word lexical categories each have only one inflectional class will still have inflectional classes. In practice, however, interest in inflectional classes only arises when a language has more than one such class for a given major lexical category. The reason for this is simple: the major question about inflectional classes is what determines membership in one or another of these classes for a given lexeme. This question is already answered for languages with only one inflectional class for a particular major lexical category, which makes the study of inflectional classes less interesting for such languages." (Aronoff 1994:182, note 6)

He thus acknowledges that there is no principled difference between one class and multiple classes; we can treat both inflection (only one class) and inflectional classes in exactly the same way.

Stump (2015:114-115) defines inflectional classes as a system of at least two classes, but his formal characterization does not really distinguish them from the canonical situation of having only one. He defines a canonical situation for inflection as presenting an isomorphy between word class, a content paradigm associated with all the members of that word class, and a form paradigm associated again with all members of that class. Inflectional classes, in his definition, only arise when some of the lexemes, although associated with the same content paradigm (set of features and values for which a given lexeme inflects), make use of a different set of realizations (form paradigm), at least partially. In fact, that definition would lead to the recognition of a separate inflectional class as soon as one form deviates for one lexeme. Interestingly enough, Stump only defines inflectional class as a property of the form paradigm for a given word class. When two

lexemes from the same word class present a different content paradigm, this is analysed as instances of defectiveness or overdifferentiation. Because the association of a given content paradigm with an entire word class is not systematic, it is not clear to me that a difference in content paradigm should not be counted as an instance of inflectional class: it will, after all, be reflected in the set of markers, as the set used for these lexemes will have a different cardinality than that of other lexemes. In addition, in cases of defectiveness, one could work under the analysis that the content paradigm is still available to defective lexemes, but that what is lacking is a subset of the realizations associated with the form paradigm. Stump's definition here is interesting because it goes against his own definition of inflectional classes as classes of stems (in Stump 2015; see below section 2.3.).

The same importance of the identity of the content paradigm between inflectional classes is noted by the canonical approach set out in Corbett (2009). Here again, the fact that inflectional classes should have the same content paradigm but a different form paradigm means that Corbett envisages the canonical inflectional class as being a non-canonical deviation from inflection proper. Thus only when there are two or more inflectional classes can we talk of inflectional class.

Some definitions allow for considering one class of inflection as an inflectional class, but all authors tend to only consider systems where there are at least two classes as being either of interest (Aronoff) or as being qualified as inflectional class systems. This is thus what will be done in this thesis. I now turn to the question of the number of lexemes required to form a class.

2.2.4.How many lexemes to form a class?

In the canonical situation as set out by Corbett (2009), a canonical inflectional class will have as close as possible a number of lexemes as its fair share in the lexicon. Thus in a system where there are 300 lexemes inflecting for a given paradigm, and three inflectional classes, canonical classes would have 100 members each. This is obviously not necessarily what is found in real life languages, where much of the time some classes have a high membership (often referred to as the 'regular classes'), and some classes have a very low membership, sometimes containing only one lexeme (the 'irregulars'). In theory, nothing distinguishes these two types of classes formally: each

lexeme still needs to make reference to a class index, or at least to a set of realizations that characterize its inflection. The only distinction that can be maintained is that between closed classes (those which do not acquire new members) and open classes. Thus irregulars are non-canonical classes with respect to their membership. The main problem in treating irregulars as a separate category is that of the threshold chosen to form a class, which will always be arbitrary: do we need two lexemes to form a class? Or three, or five? It is thus better to treat irregulars as a non-canonical instance of inflectional class.

That said, a lot of work has been produced on the processes of irregularization, which will mostly affect very high frequency lexemes (Nübling 2000; also see Bybee 2007). It is thus possible that historically the creation of new inflectional classes is different for irregular items and for classes involving a larger membership.

Similarly, a large number of linguists appeal to the concept of macro-classes (see Carstairs-McCarthy 1994; Wurzel 1987, 1989). These are a pure construct by linguists to limit the number of possible classes in a given language, much in the same way that traditional grammar limited the number of possible paradigms. Thus it is traditionally accepted that the second declension in Latin is made up of two slightly different models, one for masculine nouns, the other for neuter nouns, because of the systematic syncretism exhibited by neuter nouns between nominative and accusative. But strictly speaking, this declension is in fact composed of two separate inflectional classes, albeit non-canonical ones: their non-canonicity lies in the fact that one inherits a large number of realizations from the other. But in an inheritance hierarchy model they would still be represented as two separate classes. The same applies for the treatment of non-affixal phenomena, in particular by Carstairs-McCarthy: he considers that only those parts of the exponence that are affixal take part in inflectional classes. He thus considers that the German nouns *Tag*, pl. *Tage* 'day', and *Gast*, pl. *Gäste* 'guest' belong to the same inflectional class, which is a generalization only over their affixal marking. Strictly speaking, because of the vowel alternation on their stem, these two lexemes should be considered as belonging to two separate, albeit related, inflectional classes. In this thesis,

I thus consider inflectional classes as being global phenomena involving the whole exponence, and even slight deviations between classes as forming different classes (see Stump 2015 for a similar approach). Instead of relying on a concept of macro-classes, I model the relationship between classes in an inheritance hierarchy model (Brown & Hippisley 2012). In such a model, the inflectional classes of the language are modelled as a tree, where each node inherits a number of realizations from the preceding node, and where each node represents a separate inflectional class. Thus in Latin there would be a higher node for second declension masculine nouns of the type *dominus*, and at least two sister nodes, one for neuters of the type *templum* which would inherit most forms except the nominative and accusative, and another node for nouns of the type *ager* with a different nominative singular. What corresponds to macro-classes is modelled by the relative distance between classes in the tree. These three inflectional classes of Latin form a group which share a majority of realizations, with *ager* being closer to *dominus* because the oerride in realizations only targets one cell. The model takes into account both the differences and the similarities in the realizations shared by different inflectional classes.

2.3. Inflectional classes and Stem classes

Inflectional classes, as defined up to this point in this chapter, are classes of lexemes. From a canonical perspective, each lexeme must belong to one and only one inflectional class. A number of deviations from this canonical ideal have been defined, including heteroclisis (when a lexeme belongs to one inflectional class for part of its paradigm and to another inflectional class for another part of its paradigm). Most definitions, including that given in section 1.2 of the present work, require that each lexeme be linked to only one inflectional class, which in turn is linked to the form paradigm of the word class of a lexeme: an inflectional class is said to cover the full set of the realizations of a set of lexemes. One of the problems with this approach is that, in fact, heteroclites should be considered to form a class on their own, which would conceal a potential generalization

(the exponents are taken from two or more existing inflectional classes), but has the advantage of keeping in line with the rule of lexemes belonging to only one inflectional class. Such an analysis is not available when a given lexeme belongs to two inflectional classes for the same cells (overabundance), as is the case for at least part of the paradigm of French *asseoir, s'asseoir* 'sit'. The assumption that inflectional classes are classes of lexemes receives general acceptance in the literature (Aronoff 1994; Corbett 2009; Spencer 2013 among others; see section 1.2).

2.3.1 Inflectional classes as classes of stems (Stump 2016)

Recently, though, Stump has proposed that inflectional classes are classes of stems, not classes of lexemes (Stump 2016:93–95). His evidence comes from the existence of segregated inflectional classes, of the patterns of inflectional class distribution in heteroclisis, and of the patterns noted for overabundance. His theory also has interesting implications for phenomena such as overdifferentiation and defectiveness.

A first theoretical point of importance made by Stump concerns the status of stems in the lexical representation of a lexeme. As he points out, lexemes are purely abstract representations, and must be accompanied by a phonological representation of their lexical material, for him in the form of stem(s), which are sensitive to phonological conditioning (Stump 2016:92). This fact means that in composing rules of exponence for a given lexeme, reference to its stem, in addition to an inflectional class, is unavoidable. Stump thus argues that a rule of exponence making direct reference to a stem, rather than a lexeme linked to stems, is of a simpler format, and is thus preferable. In such a system, nothing prevents stems from being homophonous, while belonging to two different inflectional classes, as is the case for different lexemes with homophonous stems (e.g. *ring* 'put or make a ring around', past form *ringed* vs. 'sound or cause to sound', past form *rang*, Stump 2016:93), or for cases of overabundance.

Stump develops the notion of segregated inflectional classes, which for him is evidence that inflectional classes have to be considered as classes of stems. Segregated inflectional classes are inflectional classes which "only determine the exponence of a proper subset of the cells in a

paradigm" (Stump 2016:90), as opposed to the usual understanding of inflectional classes as global, that is classes whose membership "determines the realization of every cell of a member's paradigm (whether or not that realization is in all cases diagnostic of class membership)" (Stump 2016:90). He takes the example of Latin, in which, according to his analysis, any given lexeme belonging to an inflectional class in its imperfective forms need not belong to the same class in its perfective forms. One should note here that Stump includes the rules for stem formation in the computation of an inflectional class, in an approach that does not distinguish between types of exponents for given features. He thus distinguishes four conjugations for the imperfective forms, each characterized by the form of its thematic vowel: first class ā-stems, second class ē-stems, third class Ø-stems (no thematic vowel)[11], a subclass of the third conjugation for i-stems (with 1SG.PRES *-iō*), and finally a fourth class of ī-stems. What Stump notes is that one can distinguish five main rules for deriving perfective stems from the lexeme's root: -u suffixation, the bare root, -v suffixation, -s suffixation and root reduplication. He further adds that one can find examples of each of these rules for stem formation in most of the four imperfective conjugations, as indicated in Table 15 below reproducing his table (Stump 2016:91). This distinction of inflectional classes in the *perfectum* only make sense if one considers classes as global in the sense that all changes in realization are included in the computation of an inflectional class, including stem changes: for the *perfectum*, affixal realizations do not change in between classes.

[11] Stump's qualification of these stems as zero stems is somewhat misleading. It is true that none of these presents a thematic vowel, but some verbs present changes in the vowel quantity of the stem vowel, as is the case for his example *iu:ui*. A real example of a zero stem would be *uerti*.

		Imperfective			Perfective conjugation and stem				
Verbal lexeme		Conj.	Stem		Root+-*u*	Root+Ø	Impf stem + -*v*	Root + -*s*	reduplicated root
CREPĀRE IUVĀRE LAUDĀRE	'rattle' 'help' 'praise'	I	crepā- iuvā- laudā-		crepu-	iūv-	laudāv-		
MONĒRE VIDĒRE DĒLĒRE LŪGĒRE SPONDĒRE	'warn' 'see' 'destroy' 'mourn' 'pledge'	II	monē- vidē- dēlē- lūgē- spondē-		monu-	vīd-	dēlēv-	lūx- /lu:ks/	spopond-
ALERE DĒCERNERE DŪCERE CADERE	'nourish' 'decide' 'lead' 'fall'	III	al- dēcern- dūc- cad-		alu-		dēcrēv-	dūx-	cecid-
CAPERE	'take'	III(-*iō*)	capi-			cēp-			
SALĪRE VENĪRE AUDĪRE VINCĪRE	'jump' 'come' 'hear' 'bind'	IV	salī- venī- audī- vincī-		salu-	vēn-	audīv-	vinx-	

Table 15. Imperfective and perfective stems in Latin (Stump 2016:91). Shading added.

One should note that in Table 15, some of the stem forming rules correlate with the imperfective conjugation: some of the segregated inflectional classes for the imperfective correspond exactly to one class for the perfective. Thus class III (*-iō*) is only correlated with one rule for perfective stem formation. This class could thus easily be considered as global. For a number of verbs, the perfective thematic element is the only element distinguishing the inflectional realization rules between lexemes; for Stump, this thematic element is a marker of different inflectional classes, although the suffixal material realization of the feature value sets for perfective forms is identical between those classes:

			I 'prepare'	II 'remind'	III 'rule'	III (*-iō*) 'take'	IV 'hear'
Imperfect	Present	1sg	parō	moneō	regō	capiō	audiō
		2sg	parās	monēs	regis	capis	audīs
		3sg	parat	monet	regit	capit	audit
	Future	1sg	parābō	monēbō	regam	capiam	audiam
		2sg	parābis	monēbis	regēs	capiēs	audiēs
		3sg	parābit	monēbit	reget	capiet	audiet
	Preterite	1sg	parābam	monēbam	regēbam	capiēbam	audiēbam
		2sg	parābās	monēbās	regēbās	capiēbās	audiēbās
		3sg	parābat	monēbat	regēbat	capiēbat	audiēbat
Perfect	Present	1sg	parāvī	monuī	rēxī	cēpī	audīvī
		2sg	parāvistī	monuistī	rēxistī	cēpistī	audīvistī
		3sg	parāvit	monuit	rēxit	cēpit	audīvit
	Future	1sg	parāverō	monuerō	rēxerō	cēperō	audīverō
		2sg	parāveris	monueris	rēxeris	cēperis	audīveris
		3sg	parāverit	monuerit	rēxerit	cēperit	audīverit
	Preterite	1sg	parāveram	monueram	rēxeram	cēperam	audīveram
		2sg	parāverās	monuerās	rēxerās	cēperās	audīverās
		3sg	parāverat	monuerat	rēxerat	cēperat	audīverat

Table 16. Active indicative inflection of five Latin verbs (Stump 2016:77)

In the exemplar verbs given in Table 16, except for the form of the perfective stem, inflectional realizations marking all features but the feature {ASPECT: perf} are strictly identical between classes. For Stump, this is an indication that one need not specify global inflectional classes for these five lexemes, but that, for the perfective conjugation, their stem will select one of the stem forming rules, and an identical affixal realization for the feature value pairs associated with {ASPECT:perf}. It is not clear how this simplifies the formalization of classes if one considers the stem formation rules to be a sign of an inflectional class. For the imperfective conjugation, there is a thematic element, but this thematic element (a stem formation rule for imperfective stems adding a thematic vowel or nothing to the root) further correlates with at least some distinctions in the affixal realization of the associated feature value pairs. Thus a verb in class I makes use of the same affixal realizations as a verb in class II (once the stem forming element is factored out), but these select quite different realizations from the other classes. If stems have to be specified in the lexical representation, this can be done by linking some roots to specific stem formation rules. That these stems may then be linked to inflectional classes stem by stem, or globally, does not make so much of a difference: for the examples in Table 16, there would be five global classes, in addition to the

stem forming rules (each stem being represented in the lexical representation as a phonological instantiation of the lexeme linked to specific areas of the paradigm); but it would amount to six if one were to consider the perfective inflection as a separate class. In an inheritance hierarchy model, the affixal marking of feature value pairs would simply be inherited by all classes for the perfective.

The second element adduced by Stump in favour of considering inflectional classes as classes of stems is overabundance. A cell is said to be overabundant if two (or more) realizations are possible for that cell (Thornton 2012; Thornton 2013). In some cases, only one cell will be affected, as in the Italian case of the verb *seppellire* 'bury', which can have two different forms for its past participle, either *sepolto* or *seppellito* (Cappellaro 2013). In such a case, Stump considers that two conjugations are in competition, one of them being productive and with a large number of members, and the most recent form historically (*seppellito*). Each of these is attached to a different stem, so that the content paradigm will be <SEPPELLIRE, pst.ptcp>, where two stems are possible in the form paradigm <*sepol* | *seppel*, pst.ptcp>, each of them selecting a different inflectional class. Stump states that those two stems need not be phonologically different. In fact, overabundance is close to heteroclisis (see below) in that it also requires a lexeme to make reference to at least two inflectional classes, proving again for him that the global conception of inflectional classes is difficult to maintain. Such hesitation between inflectional classes is sometimes quite widespread, as is the case for Portuguese past participles (Thomas 2018). In other cases, as Stump mentions, an entire inflectional class is characterized by the fact that a section of its paradigm presents overabundance: this is the case in Spanish (Stump 2016:150–151) where overabundance is a nonexceptional property of verbs, for which there are two past subjunctive series of forms. Stump argues that this case is simply a specification of that inflectional class to supply two alternative forms rather than one (Stump 2016:151)[12].

Heteroclisis has received much attention in the past ten years (e.g. Stump 2006; Maiden 2009; Karatsareas 2011; Esher 2012:178–195; Kaye 2015; Bach & Esher 2015). Heteroclisis is the

12 It is difficult to accept Stump's argument in the case of Spanish because the two series of forms are not wholly functionally identical.

situation in which a given lexeme is linked to two or more inflectional classes. For some parts of its content paradigm forms are taken from one class, for others, from another inflectional class. Heteroclisis can involve more than two classes, and does not prevent a heteroclite lexeme from also presenting distinctive inflections for parts of its paradigm (Kaye 2015). Heteroclisis also creates a partition in the paradigm of a lexeme, in that only part of the cells will be associated with a given inflectional class. It has been noted that heteroclite paradigms tend to align the repartition of inflectional classes along stem alternation partitions of the paradigm, i.e. morphomes in the sense of Maiden (2005, 2009) and O'Neill (2011), in those languages which present such alternations. Dialectal evidence in Occitan, for example, suggests that for 'class 3b' verbs in the Languedoc, whatever the distribution of classes is in those heteroclite paradigms, it always aligns with groups of cells corresponding to a root allomorphy distribution, i.e. a morphomic structure (Esher 2012; Bach & Esher 2015). This kind of observation forms the basis for Stump's claim that inflectional classes are classes of stems: because in many cases, inflectional classes align with stem distributions in heteroclisis, it is easy to relate the two of them in a systematic way.

To some extent, Stump's theory could also account for cases of defectiveness and overdifferentiation. In the case of defectiveness for example, where the absence of forms often aligns with stem distribution patterns attested elsewhere in the language (Sims 2015), one could stipulate that a given stem is linked with no inflectional class.

2.3.2 Some possible counterevidence

In the simplest of cases, one where the inflectional classes of a language would be close to the canonical ideal, the idea of inflectional classes as classes of stems works perfectly, although in such a case there is a one to one mapping of a lexeme with a stem: the stem can enter into the composition of realization rules, and is thus needed in the lexical representation anyway. The following table presents such an example, where each stem is paired with a full inflectional class (X stands for any stem):

	I	II
SG	Xa	Xu
PL	Xb	Xv

Table 17. A simple two class system (global classes)

If we keep the assumption that an inflectional class is a full set of realizations for all the feature value sets attached to a given word class, then it presupposes that such full sets exist in the language, out of which each stem will only select a subset of realizations. That inflectional classes are classes of stems does not mean that every stem alternant must by necessity belong to a different inflectional class. Each of the stems of a given lexeme could in theory select the same inflectional class. One of the problems here is that each of these stems must be paired not only to an inflectional class, but with the relevant subset of cells associated with that inflectional class, except if we understand an inflectional class as not covering the full range of the realizations of a lexeme. The case of Inanwatan (isolate, West Papua; Vries 2004) is interesting in that respect. In Inanwatan, verbs inflect for a range of features without there being any alternations: there is only one way verbs can inflect, and the language does not present multiple inflectional classes. A small subset of verbs still presents stem alternations, with a singular stem and a suppletive plural stem[13]:

Singular stem	Plural stem	Gloss
se-	neqe-	go
aro-	tera-	moor
uwú-	te-	sit, live, stay, be

Table 18. Suppletion in Inanwatan verbs

In Inanwatan, we are faced with two alternative analyses if we follow Stump's theory. In the first, inflection is conceived as global: there is only one set of possible realizations (i.e. only one inflectional class). Most verbs only have one stem which selects the single inflectional class (the only available set of realizations for verbs) and some verbs have two stems, each of which will select only a proper subset of the same inflectional class. This seems rather a complicated analysis for a language where there are no inflectional classes. The second possibility would be to consider

13 De Vries describes the stems as referring to the number of participants in the event (Inanwatan marks subject agreement as well), not to the number of events. It does not seem to be an occurrence of verbal number.

inflection as segregated even without alternations: there would be one set of realizations for the singular, and one set for the plural. In this case, stem alternants of those verbs which present alternations will each select their adequate inflectional subset conceived as a class. But it would then force one to recognize two, homophonous stems for all the other verbs[14].

The assumption that one must pair the inflectional class with a specific subset of cells is the assumption under which a canonical definition of heteroclisis is possible: a lexeme is heteroclite if it selects for specific subsets of its realizations the realization rules attached to two or more inflectional classes already attested elsewhere in the language. This amounts to saying that the phenomenon of heteroclisis can only exist if inflectional classes are understood as classes of lexemes (or, at any rate, not as properties of stems), or at least, if all of them are attached to the full set of feature values in their content paradigm, and with a full set of possible realizations in their form paradigm. In a very simplified form, this means positing as a minimal inflectional class structure the following type, for a very simple paradigm realizing only one feature with two values (where X stands for any stem):

	I	I/II	II
SG	Xa	Xa	Xu
PL	Xb	Xv	Xv

Table 19. A simple heteroclite system with two global classes

Stump does not seem to rule out such a case, where some inflectional classes are global (class I and II in this example). In such a system though, we would have to consider that class I lexemes are only attached to one stem, to which the class index is linked; similarly, for class II lexemes. But for those lexemes left out of these two classes, one would need to posit two different stems, albeit homophonous, each of which would be attached to one of the two inflectional classes. Such a system limits the number of inflectional classes to two, but has to rely on creating a two stem

14 One could imagine a third way to deal with the Inanwatan data. Because there are no multiple inflectional classes, one could consider that the question is pointless and stems always attach to the same set of inflectional realizations. There would still be need for specifying to which proper subset of realizations each of the stem alternants could select. Also, if the stem selection of inflectional class applies to different inflectional classes, which may be created at some point in the history of the language, one does not see why inflection proper should be lexical, but inflectional classes thematic.

system for one group of lexemes without any phonological basis. The other possibility would be to consider that inflectional classes are classes of stems only, and that they need not cover the full set of inflectional realizations for a given lexeme. In such a case, one would have to posit that all lexemes have two stems: a singular stem and a plural stem, and that in this language the two stems are homophonous. In addition to that, one would have to posit four inflectional classes, segregated, which would only be relevant to a proper subset of the paradigm: an *a*-class and a *u*-class only relevant for singular cells, and a *b*-class and a *v*-class, only relevant to plural cells. In this case, such an analysis is rather uneconomical, but follows the principles that we have seen for Latin inflectional classes. A real language that is very close to this canonical heteroclite paradigm is Biak. In Biak there are no stem alternations except in the heteroclite class for the two cells which do not take their inflectional realization rules from one of the other two classes.

Biak	Class 1		Class 2	Class 3
root pattern	CC-	CV-		V-
1sg	ya-		ya-	y-
2sg	wa-		<w>	w-
3sg	i-		<y>	d-
1du.excl	nu-		nu-	nuy-
1du.incl	ku-		ku-	kuy-
2du	mu-		mu-	muy-
3du	su-		su-	suy-
3paucal/trial	sko-		sko-	sk_-
1pl.excl	nko-		nko-	nk_-
1pl.incl	ko-		ko-	k_-
2pl	mko-		mko-	mk_-
3pl.an	si-		s-	s-
3pl.inan	na-		n-	n-

Table 20. Affixal realization of subject agreement on Biak verbs (van den Heuvel 2006;Mofu 2009)

Heteroclite paradigms are no longer a problem for considering inflectional classes as global classes if they form their own class. In an inheritance hierarchy model, the fact that they use exponents coming from two different inflectional classes can be captured easily as a generalization if we allow for them to inherit forms from two different classes. The generalization thus preserves both the identity of content paradigms for all classes and the fact that there are formal similarities

between those classes.

Another, more complicated, possible scenario shows that there is a theoretical problem with considering inflectional classes as classes of stems. It concerns the extent to which the form paradigms of inflectional classes can be comparable. If each stem selects its own inflectional class, this means that not all classes should present the full range of inflectional realizations, and thus that they are not necessarily linked to the full set of feature value pairs for which a given word class inflects. In the example above, because the classes referred to by stems were global, this was not a problem. But if we consider the next, made up, example, this would create a sort of inflectional class defectiveness: some classes would have to lack some forms/cells.

	I	I/II	I/Z	II
SG	Xa	Xa	Xa	Xu
PL	Xb	Xv	Xz	Xv

Table 21. A more complex class system: heteroclisis and inflectional class defectiveness

In Table 21, considering inflectional classes as not being global and linked to a lexeme will have the result that class I and II could still be considered global, class I/II would be a class of heteroclite lexemes taking a proper subset of their realizations from class I and class II. The class I/Z would be more problematic, as the plural form does not belong to any of the existing global classes. Such lexemes could still be said to take a proper subset of their inflectional realizations from class I (singular forms), but the plural form would then have to be considered as a class of its own, which would not cover the entire possible set of feature-values and realizations for that word class. There would then be inflectional classes presenting a different content paradigm. This is possible in theory (and closely resembles cases of overdifferentiation) but it creates a new non-canonical distinction of inflectional class defectiveness, while the lexemes inflecting partially for such a class are not defective:

	I	II	III
SG	Xa	Xu	
PL	Xb	Xv	Xz

It seems easier in such a case to treat class I/Z as a class of its own simply presenting the default

realization for the singular. An example of such a heteroclite class is given by Latin thrid declension nouns.

Finally, treating inflectional classes as classes of stems also goes against Stump's own definition of inflectional classes in other works. In Stump (2015:114-115), he states that inflectional classes only differ in their form paradigm, not in their content paradigm. To talk of inflectional classes, one needs to have lexemes presenting exactly the same content paradigm but different form paradigms. If inflectional classes are classes of stems, then the content paradigm is no longer comparable, and one is faced with a very different object, which may be due to the interaction of multiple morphomic objects.

2.3.3 Classes of stems or classes of metamorphomes?

The previous section has shown that Stump's proposal to consider inflectional classes as classes of stems runs into some problematic cases. In particular, in cases of heteroclisis without stem alternations, it forces an analysis whereby all heteroclite lexeme present more than one stem even though these stems are perfectly homophonous. This does not seem desirable.

Stems have been at the centre of developments of a morphomic theory of inflectional morphology, which posits that grammar mediates syntax and phonology through a morphological component which is independent of both syntax and phonology, the morphomic level (Aronoff 1994). This level has been shown to exist in particular in cases of root allomorphy which are not conditioned by the presence of given features or by phonology (Aronoff 1994; Maiden 2013; Maiden et al. 2011), and its psychological validity proven in particular by the behaviour of such structures in diachrony (Maiden 2005) and by (some) psycholinguistic experimental evidence (Milin et al. 2009). Stem alternation structures are often referred to as 'morphomes' in the literature (Maiden et al. 2011; Cruschina, Maiden, & Smith 2013; see in particular O'Neill 2011; O'Neill 2014). Another type of morphomic structure is inflectional classes (Aronoff 1994).

Erich Round has recently proposed a typology of these morphomic phenomena or objects, in

which he distinguishes three types, rhizomorphomes, metamorphomes and meromorphomes (Round 2015). The last type, meromorphomes will not be relevant for this study: it is a specific type of morphomic object which mediates the morphotactics of some languages such as Kayardild (Round 2013), and has until now only been identified in a handful of languages belonging to the Tangkic family. Rhizomorphomes are morphomic structures related to roots, i.e. inflectional classes (considered as classes of lexemes), and metamorphomes are morphological structures establishing a partition in the paradigm for the specific behaviour of a group of cells, what has often been equated with stem alternation patterns, or at least has been best exemplified by such alternations. Round proposes different names for these structures rather than the names which are usually used to refer to them in order to emphasize their commonality: all three types of structures are morphomes, and belong to the morphomic level of grammar.

Morphome types	Pertain to	Divide up	By similarity of
Rhizomorphomes	sets of roots	the lexicon	paradigms
Meromorphomes	sets of word formation operations	morphological mappings	patterns of exponence
Metamorphomes	sets of cells in a paradigm	paradigm types	incidence of (realizations of) meromorphomes

Table 22. Types of morphomes (Round 2015:29)

Table 22 presents the typology established by Round with some definitions. He clearly considers inflectional classes (rhizomorphomes) to be lexical. In addition, metamorphomes are understood as properties of a set of cells characterized by a specific, similar behaviour in some respect, which may, but need not be, instantiated by stem alternations. Metamorphomes are thus a group of cells creating a partition in the paradigm, aligned or not on extramorphological factors but most visible when not aligned, which governs specific morphological behaviours of the inflectional realization of that group of cells as compared to the rest of the paradigm. This definition is interesting in that it becomes possible to treat as metamorphomes a number of structures proposed recently in the literature: the stem spaces first proposed by Bonami and Boyé (2003, 2007), the morphomes understood as metamorphomes (e.g. Maiden 2005) and finally the inflectional zones

proposed to account for heteroclisis (Walther 2013; see also Maiden 2018).

Stump, by positing that inflectional classes are classes of stems (or classes of metamorphomes if one follows the analysis I propose here: but their indexing is 'morphomic') is also positing a dependency relationship between two types of morphomic objects, rhizomorphomes and metamorphomes. The relationship he establishes is directional: stems, in a way, take precedence in his theory over rhizomorphomes. The relationship between these two types of structure has been shown to be extremely complex, in particular in diachrony (Esher 2012, 2017) and there does not seem to be agreement with regard to a potential directionality of the relationship. Furthermore, by positing such a systematic relationship between the two types of structures, one risks losing some of the important generalizations which arise from their separate study. If in some languages the range of metamorphomes possible in combination with a single given inflectional class varies, in others it does not. In some cases, one could argue for a reverse directionality: thus in French, verbs belonging to the first conjugation either show no stem alternations or a very restricted set of possible alternations (Esher 2017; Maiden 2018:277-283).

Applying to French the analysis made by Stump for the Latin verb would run into a further problem. French verbs present from one to nine different stems. The number of stems does not seem to be correlated with the inflectional class chosen. In the case of a verb with nine stems such as *pouvoir* 'can', this would mean than a verb would select nine different (or possibly identical) inflectional classes. Similarly, a verb with one stem would select only one. But this means that these inflectional classes have a very different content paradigm and are thus not comparable, while it seems desirable for all lexemes to present similar inflectional structures. The alternative would be to consider a maximal number of stem spaces conceived of as partitions of the paradigm which behave in a specific manner with regard to stem selection, and consider that all lexemes present these partitions. There would be an inflectional class chosen for each stem space, which would mean that a paradigm would be made of nine stem spaces and nine inflectional classes for each verb in French. This analysis does not seem acceptable as it posits too many homophonous stems, and creates

structures which are not formally distinguished by actual lexemes in the language.

2.4. Inflectional classes and word classes: sharing inflectional classes

The question of the sharing of inflectional classes across word classes is of importance for understanding the diachronic development of inflectional classes. There is no doubt that different word classes can share similar exponents and a large number of features with an identical range of possible values. A good example can be found in the typical marking of all the members of a noun phrase with identical markers in Bantu languages[15]. Consider the data from Chichewa in Example 1 and 2:

1) chipewa chi-modzi
   hat 7-one
   'one hat'

2) zipewa zi-wiri
   hats 8-two
   'two hats' (Data from Corbett 1991:107)

In the above examples, a number of features are shared by all the members of the NP, and all bear marking for the same value: *chi-* in Example 1 minimally realizes the feature value pair {NUM:SG}.

The canonical typology of inflectional classes presented by Corbett (2009:7) makes the point that a canonical inflectional class should not be constrained by word class membership. In other words, a canonical inflectional class should not appear to be available only to verbs. The

15 Not all Bantu languages make use of absolutely identical markers for all targets for agreement. Swahili for instance sometimes makes use of different realizations on the noun, the adjective and, for subject agreement, on the verb (see Schadeberg 2001). In fact, no language with absolute alliterative concord has yet been found (Corbett p. c.), and languages at most do present some alliterative realizations. Thus in Arapesh, most singular forms may be considered as alliterative, but it does not stand for plural forms.

question remains open whether this canonical ideal might exist in real language situations, and what conditions are necessary for such a sharing to hold. Obviously, it is not necessary for the canonical ideal to exist in actual languages, but if one of the criteria for canonicity were to never hold, such a criterion could be considered as nondistinctive, and thus inoperative: if one finds that no language presents inflectional classes shared across two or more word classes, it would follow that all inflectional classes are non canonical in that respect. The question largely depends on the definition of inflectional class. Corbett, like most authors referred to here, broadly follows a definition similar to that of Aronoff (1994:64) for whom inflectional classes are a set of lexemes sharing identical realizations for an identical set of feature-value pairings. According to such a definition, an inflectional class could only be shared between word classes if both the markers and the features that can be marked on the lexemes are identical. Put into Stump's theory, an inflectional class can only be available to different word classes if these word classes share an identical content paradigm and an identical form paradigm.

Two different inflectional classes for lexemes of the same word class will normally be distinguished by exhibiting different realizations for the same set of feature value pairs. This is the canonical definition of inflectional classes where the content paradigm is constant while the form paradigm is different (see Corbett 2009:4).

Similarly, if two lexemes do not realize the same set of feature value pairs, their inflectional class will be different, even though the realizations of these feature value pairs might make use of similar or identical phonological content (the markers can be identical but realize different features). Good examples of homophonous markers are not difficult to find. In many Australian languages, at least one of the markers of ergative case on nouns was also grammaticalized as a marker of tense on verbs.

What about full inflectional classes? One could easily imagine a situation where adjectives and nouns inflect in a very similar way. Historically, this might arise from all being underspecified nominals, which then specialized in different possible syntactic positions: restriction in the syntactic

slots that these nominals can occupy might have happened without those changing their inflectional behaviour.

2.4.1 Why examples cited in the literature do not work

Most of the examples of sharing of inflectional classes between word classes in the literature do indeed come from sharing between adjectives and nouns. Famous examples include Russian (Corbett 2009) and Latin (Walther 2013). Although all these examples show striking similarities between the inflectional classes of nouns and adjectives (they exhibit identical exponents for at least a subset of the class), they all fail to show identical inflectional classes if one follows the definition given by the same authors. In all cases, two crucial criteria make these classes different in synchrony: the size of the paradigm, and the presence of an added feature for adjectives not exhibited in the inflectional system of nouns, gender.

In the example of inflectional class shared between nouns and adjectives in Russian adduced by Corbett (2009) as a case where an inflectional class is available to both adjectives and some nouns, gender is the feature that in fact makes the two inflectional classes different. In Table 23, *stolovaja* 'dining room' is a noun, while *staraja* 'old' only shares with it the section of its paradigm characterized by the value feminine for the feature gender. The fact that the noun *stolovaja* may be feminine does not mean that it is inflected for a feature GENDER, and thus, the content paradigm of the two classes is different. The feature gender in a noun is of a very different nature than it is in an adjective: it is a morphosemantic feature in a noun, and a morphosyntactic feature in an adjective. The possible range of values in the paradigm is thus different: the paradigm of an adjective exhibits all the possible values for the feature gender, while a noun is restricted to only one value. Thus, *stolovaja* only shares part of the paradigm available to adjectives, lacking the forms inflecting for masculine and neuter.

	komnata 'room'	stolovaja 'dining room'	staraja 'old' (F SG)
NOMINATIVE	komnata	stolovaja	staraja
ACCUSATIVE	komnatu	stolovuju	staruju
GENITIVE	komnaty	stolovoj	staroj
DATIVE	komnate	stolovoj	staroj
INSTRUMENTAL	komnatoj	stolovoj	staroj
LOCATIVE	komnate	stolovoj	staroj

Table 23. Potential sharing of inflectional class between some nouns and adjectives in Russian (from Corbett 2009:7)

Another example often cited for inflectional classes available to multiple word classes is Latin nouns and adjectives. Table 24 presents the inflectional realizations of lexemes belonging to what is traditionally labelled the first and second declension, consisting in fact of three main inflectional classes. Table 25 presents the realization of the first class adjective *bonus* 'good' (vowel quantity is only indicated where it is relevant).

	Second declension		First declension
	acinus 'berry'	*templum* 'temple'	*rosa* 'rose'
SG.NOM	acinus	templum	rosa
SG.VOC	acine	templum	rosa
SG.ACC	acinum	templum	rosam
SG.GEN	acini	templi	rosae
SG.DAT	acino	templo	rosae
SG.ABL	acino	templo	rosā
PL.NOM	acini	templa	rosae
PL.VOC	acini	templa	rosae
PL.ACC	acinos	templa	rosas
PL.GEN	acinorum	templorum	rosarum
PL.DAT	acinis	templis	rosis
PL.ABL	acinis	templis	rosis

Table 24. First and second declension nouns in Latin

	Masculine	Neuter	Feminine
SG.NOM	bonus	bonum	bona
SG.VOC	bone	bonum	bona
SG.ACC	bonum	bonum	bonam
SG.GEN	boni	boni	bonae
SG.DAT	bono	bono	bonae
SG.ABL	bono	bono	bonā
PL.NOM	boni	bona	bonae
PL.VOC	boni	bona	bonae
PL.ACC	bonos	bona	bonas
PL.GEN	bonorum	bonorum	bonarum
PL.DAT	bonis	bonis	bonis
PL.ABL	bonis	bonis	bonis

Table 25. A first declension adjective in Latin, bonus *'good'*

As in the Russian example above, exponents are identical. The section of the adjectival paradigm marked for feminine gender parallels realizations for first class nouns, the sections marked for masculine and neuter gender, those of the two models for second class nouns. The problem here is that they only share realizations and most of the features realized, not all of them[16]. Again, the fact that nouns in these two classes may be feminine, masculine, or neuter does not account for the similarity in synchronic terms, because for nouns gender is an inherent feature, for which they do not inflect, and for adjective it is a contextual feature of agreement with the noun head of the NP. In addition, some of the members of the first class of nouns are masculine (*agricola* 'farmer', *poeta* 'poet', etc), and so do not even share a similar value in gender with the feminine part of the adjectival paradigm. Similar to the Russian example is the fact that these nouns would only share part of the paradigm available to adjectives, not a full paradigm. If one considers an inflectional class to be a class of lexemes, it is thus impossible to consider that adjectives and nouns in Latin can belong to the same inflectional class.

16 That adjectives should share common features with nouns in Latin can be easily accounted for because there is agreement for number and case between nouns and adjectives. There is gender agreement as well, but nouns do not inflect for gender.

2.4.2 Looking for good candidates for sharing of inflectional classes across word classes

It seems unlikely that nouns and adjectives would be good candidates for the sharing of inflectional classes, because of the discrepancy in the presence of a gender feature in their inflectional system, except in languages with no gender, such as Finnish. Good candidates should be sought in word classes that share an identical set of feature value pairs. The preceding section has shown that for two word classes to share an inflectional class, the two word classes must mark exactly the same set of features inflectionally. The fact that verbs and nouns typically mark very different features in inflection (TAM features for verbs, number and case for nouns) makes it unlikely that a good candidate would be found in the inflectional classes of nouns and verbs. Some languages do mark TAM features on nouns (see Nordlinger & Sadler 2004), but most of these also mark other features either on verbs or on nouns. A more likely scenario will involve small, closed word classes known to regularly evolve from larger, open classes (Heine and Kuteva 2007). Evolutionary scenarios, particularly in grammaticalization theory, often develop the idea that at the times of glossogeny, only word classes referring to events and entities are available (properties can be referred to with either of those, as is the case in many languages without a specific class of adjectives, or with a very restricted class of adjectives; such a situation is common in Austronesian and Pama-Nyungan languages). Thus, Heine and Kuteva (2007) compile evidence and come up with possible scenarios for the emergence of new word classes, as shown in Figure 1:

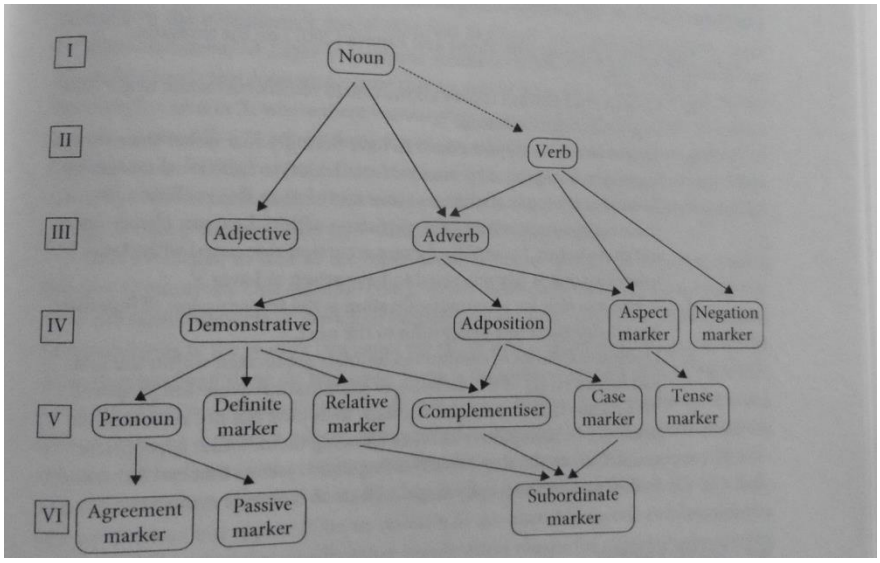


Figure 1. Layers of grammatical evolution (Heine & Kuteva 2007)

The most probable candidates come from two word classes, one of which at least is a closed, smaller word class, that could evolve as a different word class from an already inflecting word class, and keep its inflection. Such examples might well involve mostly minor, closed word classes, with very few members. A possible candidate would be the case of the possessive pronouns/adjectives in Latin, which share a complete paradigm with first class adjectives. Possessive pronouns/adjectives can be distinguished from adjectives as a word class in that they can be used pronominally, which is not the case with adjectives. The distinction is thus made on syntactic (distributional) grounds. But one should note that identity of inflection might also have been kept because in most occurrences, these pronoun/adjectives are used adjectivally: the distinction of these word classes is small, and much is kept in common between them. Table 26 presents the paradigm for the second person singular possessor pronoun *tuus*, and Table 27 an exemplar paradigm for first class adjectives:

	Masculine	Feminine	Neuter
SG.NOM	tuus	tua	tuum
SG.VOC	tue	tua	tuum
SG.ACC	tuum	tuam	tuum
SG.GEN	tui	tuae	tui
SG.DAT	tuo	tuae	tuo
SG.ABL	tuo	tuā	tuo
PL.NOM	tui	tuae	tua
PL.VOC	tui	tuae	tua
PL.ACC	tuos	tuas	tua
PL.GEN	tuorum	tuarum	tuorum
PL.DAT	tuis	tuis	tuis
PL.ABL	tuis	tuis	tuis

Table 26. The 2SG possessor possessive pronoun/adjective in Latin

	Masculine	Feminine	Neuter
SG.NOM	bonus	bona	bonum
SG.VOC	bone	bona	bonum
SG.ACC	bonum	bonam	bonum
SG.GEN	boni	bonae	boni
SG.DAT	bono	bonae	bono
SG.ABL	bono	bonā	bono
PL.NOM	boni	bonae	bona
PL.VOC	boni	bonae	bona
PL.ACC	bonos	bonas	bona
PL.GEN	bonorum	bonarum	bonorum
PL.DAT	bonis	bonis	bonis
PL.ABL	bonis	bonis	bonis

Table 27. A first class adjective in Latin, bonus *'good'*

Clearly, the two paradigms present identical realizations for identical features, as can be seen in Table 28:

	Masculine	Feminine	Neuter
SG.NOM	-us	-a	-um
SG.VOC	-e	-a	-um
SG.ACC	-um	-am	-um
SG.GEN	-i	-ae	-i
SG.DAT	-o	-ae	-o
SG.ABL	-o	-ā	-o
PL.NOM	-i	-ae	-a
PL.VOC	-i	-ae	-a
PL.ACC	-os	-as	-a
PL.GEN	-orum	-arum	-orum
PL.DAT	-is	-is	-is
PL.ABL	-is	-is	-is

Table 28. Exponents of first class adjectives and possessive pronoun/adjectives in Latin

The realization rules common to the two paradigms can be expressed simply following Stump's model with a content paradigm and a form paradigm. Both are identical for both *tuus* and *bonus*. Thus for nominative singular, the following set of rules can be formalized:

Content	Form	Realization
<BONUS, {nom sg masc}>	<*bon*-, {nom sg masc}> where PF= <X {nom sg masc}> --> X-*us*	bonus
<BONUS, {nom sg fem}>	<*bon*-, {nom sg masc}> where PF= <X {nom sg masc}> --> X-*a*	bona
<BONUS, {nom sg neut}>	<*bon*-, {nom sg masc}> where PF= <X {nom sg masc}> --> X-*um*	bonum

Table 29. PFM rules for Latin adjectives

Content	Form	Realization
<TUUS, {nom sg masc}>	<*tu*-, {nom sg masc}> where PF= <X {nom sg masc}> --> X-*us*	tuus
<TUUS, {nom sg fem}>	<*tu*-, {nom sg masc}> where PF= <X {nom sg masc}> --> X-*a*	tua
<TUUS, {nom sg neut}>	<*tu*-, {nom sg masc}> where PF= <X {nom sg masc}> --> X-*um*	tuum

Table 30. PFM rules for the Latin possessive tuus *'your'*

Are these perfect examples? One might argue that the whole system of these pronouns in fact exhibits a very different paradigm, including features of person and number for the possessor. This would make the possessive pronoun a single lexical item, not a closed word class comprising five members *meus* 'my/mine', *tuus* 'your / yours', *suus* 'his / her / hers / their / theirs', *noster* 'our / ours', *vester* 'your / yours (PL)'. If one is to separate these words into five different lexical items, the underspecification of *suus* for possessor number is more difficult to account for. In the single lexeme account, it can simply be accounted for with syncretism, much in the same way as *your* would be in English. The question of the sharing of inflectional classes between pronouns and other word classes hinges on the definition of lexeme to a great extent. If the possessive adjective / pronoun is considered a single lexeme, the identity of inflectional class with adjectives fails, as the possessive also realizes features for possessor number and person: the form of the paradigm is different, and the possessive pronoun/adjective realizes a larger number of features (possessor person and number) if considered to be a single lexeme, as shown below in a reduced form for

nominative only[17]:

Possessor → Possessed ↓	1SG	2SG	3SG	1PL	2PL	3PL
3SG	meus	tuus	suus	noster	vester	suus
3PL	mei	tui	sui	nostri	vestri	sui

Table 31. Latin possessives as a single paradigm

One of these items also presents a further problem: strictly speaking, *meus* follows a different pattern of inflection that one could consider a subclass, in that it presents a different realization rule for its vocative singular masculine (*mi*, historically a possessive dative instead of the expected ***mee*). Additionally, *tuus* stands out by being defective: there is no form for the vocative.

Other possible candidates for the sharing of inflectional class involve two, very small, closed classes of words: ordinal numerals and indefinite pronouns / adjectives. Thus the ordinal numeral *unus* 'first' shares both a form paradigm and a content paradigm with the indefinite pronoun / adjectives *solus* 'alone', *totus* 'all' and *nullus* 'none'. One should still note that while the indefinites only inflect for singular number, *unus* also inflects for plural number when used with lexical plurals such as *castra* 'military camp'.

2.4.3 Extending principles for the sharing of inflectional classes across word classes

From the analysis of the various candidates for the sharing of an inflectional class across word classes, it is possible to devise general principles, which govern the possibility for two word classes to share an inflectional class. These principles have not featured in the literature and are a new addition made to clarify the definition of inflectional class. These principles work in the reverse way as well, in order to predict cases in which such a sharing would be impossible. This has implications for the analysis of inflectional classes in diachrony. The principles are meant to

17 This point is obviously less of a problem if one is to consider inflectional classes as classes of stems (see Section 2.3.). In such a scenario, even a putative lexeme POSS.PRON would have five stems, four of which would share the inflectional class for first class adjectives as a whole (the only exception would be *meus*, which would form a subclass of it).

constrain the possibilities offered by the canonical approach to inflectional classes: the theoretical space drawn by the canonical approach is constrained by possible and impossible cases in real instances of languages. The preceding section has shown that there are cases in which an inflectional class is shared by two word classes (possessive pronouns and first class adjectives in Latin); the following new principles show that such cases are deemed to be the exception rather than the rule, but that there are indeed cases where different word classes can share an inflectional class fully.

PRINCIPLE 1: The sharing of exponents does not indicate the sharing of an inflectional class (it is a necessary but not a sufficient criterion).

In the preceding section, I have analysed examples from Russian and Latin which show that the sharing of phonological realizations does not indicate that exactly the same features are realized on the lexeme. In Russian, the noun *stolovaja* 'dining room' presents the same exponents as the feminine adjective *staraja* 'old (F)', but the noun does not realize the feature gender. Homonymy of inflectional markers is not exceptional: studies of syncretism have shown how identical forms can express various bundles of feature value pairings inside the same paradigm (e.g. Baerman et al 2005). A number of languages also happen to use similar affixes for very diverse functions, either because of the semantic extension of the range of meanings taken by a marker (as is the case for Biak *be-*), or because of sound changes having created accidental homonymy.[18] It is thus not surprising that two inflectional classes linked to two different word classes may present similar or identical realizations, although presenting a completely identical set of realizations which would realize completely different features does not seem likely (see Principle 2). But an identical set of exponents will indicate an identical inflectional class if and only if the set of features and their possible values is identical, that is, if both the content paradigm and the form paradigm are identical.

18 There is still evidence that, at the phonetic level, such cases are not fully identical. In other words, speech production is sensitive to morphological structure (Zimmermann 2016; for a diachronic correlate of these facts, Maiden 1991).

PRINCIPLE 2: The sharing of exponents across two word classes may indicate a common origin for the two word classes (one word class originally was divided into two word classes).

Inflectional classes develop from inflectional marking (see Chapter 3). Because inflection develops mostly from the phonological coalescence of grammatical markers with words having a lexical meaning, the grammatical meaning of the marker of inflectional values tends to have relevance to the word it coalesces with. In the simplest of cases, the inflectional values grammaticalized and morphologized in inflection will be the features most relevant to the word class to which a lexeme belongs: in such a case, inherent inflection, which is a matter of choice for the speaker. Examples will include TAM features for verbs, and number and case for nouns. Such features are unlikely to develop on more than one word class at a time. Subsequent developments may involve agreement features between word classes (e.g. participant marking on verbs from pronouns, as is the case in most Austronesian languages), but the source of the marking will be different. In any case, if two word classes share an identical set of features realized through inflection, and an identical set of realizations, it might indicate that these word classes were originally part of the same word class. In the Latin example of first class adjectives and possessive pronoun / adjectives, it is not difficult to imagine a diachronic scenario in which possessives first arose from an adjectival derivation of (personal) pronouns. The stem chosen for such a derivation was that of the genitive and dative cases, which were used for attributive possession. As possessive adjectives, these forms were no different from other adjectives for their inflection, and were assigned to the first class of adjectives. Subsequently, these adjectival forms may have been used in isolation through nominal ellipsis: *whose dog is that? it's my (dog)*. The fact that these forms could also be used pronominally meant that syntactically they formed a separate (albeit related) word class.

The development of auxiliaries from full lexical verbs may also reflect a similar grammaticalization path whereby auxiliaries and full lexical verbs may share an inflectional class while being ascribed to two, related word classes (e.g. lexical and auxiliary *have* in English).

PRINCIPLE 3: In the presence of a gender system, no inflectional class available to nouns can be shared with the word classes targeted for gender agreement.

All the examples adduced until now in the literature for cases where two word classes share an inflectional class involved nouns and adjectives (in Russian and Latin). I have shown above that in the presence of a gender system, nouns and adjectives agreeing for gender cannot share an identical inflectional class, because their content paradigm is different: the classes may be related (they share a number of features, and identical realizations), but they cannot be considered to be the same class. Principle 3 extends this analysis towards all targets of gender of a gender system. In presence of a gender system, all targets for gender agreement will minimally realize the feature gender in one of its values, whatever the other features realized may be. Because nouns (the controller of gender) do not realize gender, they cannot share an inflectional class with the targets for gender agreement.

One possible counter-example comes from Latin, which presents the nouns *amicus* 'male friend', and *amica* 'female friend', as well as a first class adjective *amicus* declining for the three gender agreement values masculine, feminine and neuter. The fact that it is the same word, sharing realizations between *amicus* N.MASC and the masculine of the adjective, and *amica* N.FEM and the feminine forms of the adjective could lead to think that these are sharing an inflectional class. But in any case, the feature gender for the nouns and the adjective is still of a different nature: in one case it controls agreement and only has one value for any given nouns, being used for derivation in the feminine and masculine much in the same guise as doublets such as *dominus* and *domina*, with no corresponding adjective. In addition, there is no nominal form ***amicum* as a neuter noun.

One might be tempted to extract a general tendency from the example of actual sharing of inflectional class presented above for Latin between possessive pronoun/adjectives and first class adjectives. Both are targets for gender agreement –in fact, targets for agreement with a noun controller for the features gender and number. It might well be that in a system presenting gender,

cases of inflectional class sharing across word classes is to be found within those word classes which are targets for gender agreement. In the case of Latin, this restricts the possibilities to the word classes that can appear in a noun phrase.

PRINCIPLE 4: The sharing of inflectional classes will not be possible between controllers and targets for agreement.

This principle is a tentative extension of Principle 3 to all systems of agreement. If the controller and targets for gender agreement cannot share an inflectional class, it is possible that in any system of agreement, the controller and targets cannot share an inflectional class either. This may have to do with the distinction between inherent and contextual inflection: if marked at all for a feature entering in a system of agreement, the controller will mark it as inherent inflection. This means that the marking of the feature is a matter of choice for the speaker. It also means that the controller only presents one of the possible values that the targets could inflect for, and thus a different content paradigm. In the case of agreement in number between a noun and a determiner, for example, the speaker choses the form of the noun to refer to either one or more entities. The marking of number on the determiner in such a case, on the contrary, is not a matter of choice, but follows from the fact that the noun is marked for number. In the case of a gender system, the noun may not be marked for gender inflectionally, but such a situation is by no means restricted to gender. In French, most nouns are not marked for number.[19] Number does not appear directly on the noun, but only through agreement on the determiner, and subject-verb agreement:

3) la petite fille boit

   [la ptit fij bwa]

   'the small girl drinks'

4) les petites filles boivent

19 The only class of nouns which is marked for number in Modern French is the alternating *-al* or *-ail* SG / *-aux* PL class, such as *un vitrail* [vitʁaj] / *des vitraux* [vitʁo]. Some other nouns mark number through various types of realizations, all of which are idiosyncratic to one or two lexemes (e.g. *bœuf, œil*).

[le ptit fij bwav]

'the small girls drink'

In Example 3 and 4, the noun *fille* [fij] does not inflect for number. In Example 4, the form [fij] refers to more than one entity, which is only realized through agreement on the determiner [le], and on the verb [bwav]. The targets for agreement realize number as a contextual, morphosyntactic feature.

The same may apply for subject agreement on verbs for languages where verbs do not realize any other features, such as South Halmahera-West New Guinea languages. In most of these languages, verbs do not mark TAM features, but only inflect for subject agreement. Is it possible in such a case to say that verbs and pronouns may share an inflectional class? Pronouns and verbs do not exactly share realizations: bound pronominal forms are marked on the verb. There is also a difference in the fact that for the subject of a clause, which may not be marked for person in case of an NP, the inflectional marking of the pronoun is inherent inflection, while it is contextual in verbs. This might also explain why, in most cases, the agreement markers take a slightly different form than the full pronoun.

2.5. Conclusion

Inflectional classes can be defined as a deviation from canonical inflection. They are sets of lexemes, not sets of stems, whose members all select the same set of inflectional realizations for a similar set of feature value pairings. This point is important: two inflectional classes will be in a system of inflectional classes only if their content paradigm is identical. Inflectional classes are thus fully comparable. The distinctions in terms of the form paradigm can be minimal: if only one form is different we are presented with a different inflectional class, although one that inherits most of its exponents from another inflectional class. In such a case, the class will be non-canonical in that it does not exhibit full formal differentiation from other classes. The number of lexemes does not

matter to form a class: in such a system, irregulars are simply inflectional classes with only one member, which do not differ formally from classes with higher membership. Most often such classes with a low number of members will exhibit high token frequency: the lexemes which belong to such a class are of high frequency in a corpus. The points developed in this chapter have aimed at providing a clearer definition of inflectional classes, which is the definition used throughout this work. One can now turn to the diachronic emergence of inflectional classes.

3. The possible origins of inflectional classes: an overview

This chapter gives an overview of all the possible origins for inflectional classes that I could find either in the literature or by analysing extant systems. It makes use of the definition of inflectional classes provided in the preceding chapter and only takes into consideration classes conforming to that definition. The following chapters will provide further discussion of some of the possible origins for inflectional classes referred to in this chapter.

The principle underlying this analysis is that, in theory, any kind of formal, systematic alternation between lexemes, co-occurring with inflectional marking in word forms, could give rise to inflectional classes. For example, a number of derivational affixes deriving verbs from nouns could co-occur with inflectional marking. If their derivational status ceased to be apparent, through the progressive coalescence of inflectional affixes with the derivational ones, inflection class distinctions would arise between lexemes formerly presenting distinct derivational affixes A and B. This means that in order to find out what can give rise to inflectional classes, one has to determine the various ways in which an alternation between lexical classes can arise: anything that can create at least two groups of lexemes in a given word class that behave differently, and whose behaviour can have an influence on the way inflection is formally realized. In any case, in order for a language to acquire inflectional classes, it must first present inflection of some sort.[20] I will not study here the ways in which a language can acquire inflection (potentially through grammaticalization) but assume that those languages first acquired an inflectional system.[21]

In her comprehensive study of conjugation classes in the Germanic languages, Antje Dammel

20 There is in theory one case where a language could acquire inflectional classes without previously presenting inflection: if a language morphologized some inflectional distinction for only a subset of lexemes of a lexical category X, where one would consider inflecting and non inflecting lexemes as members of two different inflectional classes.

21 There is another case where one can assume that a language acquires both inflection and inflectional classes at the same time. In the case of the grammaticalization of some inflectional feature, there will be sound adaptation occurring between the new affix and the stem, which can in certain conditions create allomorphy, and thus inflectional classes. This is for example what happened in SHWNG languages when subject agreement markers grammaticalized: two classes arose immediately depending on the fact that the stem started in a consonant or in a vowel. See Chapter 4 on this precise case. In all other possible scenarios, the language must first have acquired inflection before inflectional classes can be created.

distinguishes five possible origins for inflectional classes in those languages (Dammel 2011:268-273; also see the review in Enger 2012). These five possibilities cover the whole of the attested possibilities examined in the literature. She distinguishes sound change proper from reductive sound change conditioned by the frequency of a lexeme; categorial reinterpretation from grammaticalization; and finally reanalysis (Dammel 2011:268-273). These five major cases for inflectional classes which arise in fact largely come down to three mechanisms: mechanisms of sound change (sound change proper, reductive sound change conditioned by frequency), mechanisms of grammaticalization (categorial reinterpretation and grammaticalization proper), and reanalysis, of the type termed 'secretion' by Haspelmath (1995) which involves a distinction of the place assigned to sounds in the morphological structure, either in the stem or in the ending.

This chapter gives a first overview of the possible types of origin for inflectional classes, some of which will be developed and analysed further in the next chapters. The basic idea underlying this chapter is that any kind of alternation that can encode an inflectional value or that targets an item associated with the marking of an inflectional value can give rise to an inflectional class. That means that, in theory, most alternations present in a given language could give rise to an inflectional class. In particular, any alternation, meaningful or not, associated with a given lexical class or classes is prone to give rise to inflectional class distinctions. Some of the distinctions correspond to the possibilities assumed in Dammel (2011), to which I add the further possibilities of heteroclisis, deponency, lexical strata, and classes coming from gender systems and alienability distinctions.

This chapter comprises two parts. In the first part, I imagine what types of formal alternations could give rise to inflectional classes, and what conditions should be met for these alternations to be morphologized to the point where they only signal formal alternations. I track the origins of classes in phonological alternations fossilized inside word boundaries (i.e. word level phonology, see Inkelas 2014); in sound change applying across multiple cells of inflectional paradigms and giving rise to new alternations. I then examine the cases of reanalysis and grammaticalization. Attention is given to inflectional classes arising from lexical strata and from heteroclisis. Finally, two major

sources of inflectional class systems are discussed: alienability splits in possessive inflection, and gender systems. In the second part of the chapter, I ask what classes are impossible, examining in particular the claim that the No Blur Principle would constrain the possible systems.

3.1. Inflectional classes from sound change

Sound change is often said to be the main origin for inflectional classes, sometimes the sole scenario presented by authors (e.g. Carstairs-McCarthy 2010; see Chapter 4). There are at least four possible scenarios for inflectional classes to originate from phonological alternations or sound contrasts. One can imagine inflectional classes arising from sound change, from the change of phonological rules, and finally from reanalysis and the loss of the conditioning environment for a sound alternation. A further scenario has to do with the phonological reduction of lexemes dues to their frequency.

3.1.1 Sound change

The first possibility applies to sound change proper, applying to inflected forms, in particular at affix boundaries. This is a case where one class of inflection gives rise to multiple classes. Here is a possible scenario, given for the sake of clarity in an invented language.

Let us imagine a language presenting the following paradigm of two verbs, ABA and ERD, in the past and non-past tenses, for the six person and number combinations, as shown in Table 32 and 33:

ABA		1	2	3
Past	SG	aba-da	aba-di	aba-du
	PL	aba-ri	aba-gi	aba-gu
Non-past	SG	aba-ga-da	aba-ga-di	aba-ga-du
	PL	aba-ga-ri	aba-ga-gi	aba-ga-gu

Table 32. Made-up paradigm for ABA

ERD		1	2	3
Past	SG	erd-da	erd-di	erd-du
	PL	erd-ri	erd-gi	erd-gu
Non-past	SG	erd-ga-da	erd-ga-di	erd-ga-du
	PL	erd-ga-ri	erd-ga-gi	erd-ga-gu

Table 33. Made-up paradigm for ERD

In such a language, although person and number are marked cumulatively with a single affix, the canonical inflectional ideal of a one-to-one mapping is close to being met. In addition to that, there is only one class of verbs, all marking the same inflectional distinctions in a formally identical way. Past is unmarked, and-non-past is marked closer to the root than the person and number agreement for subject. Now, let us imagine that a sound change arises in the language, that specifically targets voiced stops between vowels and yields the following results, attested in many languages: d>z, g>Ø, b>v. The new paradigms thus created would look as shown in Table 34 and 35, provided that this sound change applied regularly and without exceptions (some possible outputs of the cooccurrence of sequences of vowels are added, positing that hiatus is impossible, that length is phonemicized, and that syllable nuclei cannot contain more than two morae):

AVA		1	2	3
Past	SG	ava-za	ava-zi	ava-zu
	PL	ava-ri	ava-i>avaj	ava-u>avaw
Non-past	SG	ava-a-za >ava:za	ava-a-zi >ava:zi	ava-a-zu >ava:zu
	PL	ava-a-ri>ava:ri	ava-a-i>avaj	ava-a-u>avaw

Table 34. Sound change applied to the made-up paradigm AVA

ERD		1	2	3
Past	SG	erd-da	erd-di	erd-du
	PL	erd-ri	erd-gi	erd-gu
Non-past	SG	erd-ga-za	erd-ga-zi	erd-ga-zu
	PL	erd-ga-ri	erd-ga-i >erdgaj	erd-ga-u >erdgaw

Table 35. Sound change applied to the made-up paradigm ERD

The result is easily formalized as the creation of two inflectional classes. Some allomorphy for the past singular cells can be accounted for in terms of phonology (alternation of contexts for stop lenition). It might be possible to suggest a similar explanation for past 2/3PL, except that an

unexpected result of the sound change is an accidental syncretism between past and non-past 2/3PL for the AVA-class. One should note that, except for this syncretism, the sound change leaves a system where all the contrasts that were previously realized are still realized contrastively. Thus non-past would be realized affixally in the ERD-class, but through stem vowel lengthening for the AVA-class. Such an alternation for the past paradigm is impossible to account for in terms solely of the phonology of the language: this is a clear instance of the creation of inflectional classes through sound change. In this case, the sound change has applied mainly on the affixes, but it is of course possible that it targets the final segment of the stem (one possible origin for thematic markers in systems where one class lacks a thematic marker). The result of sound change gives the two inflectional classes shown in Table 36 and 37.

AVA		1	2	3
Past	SG	-za	-zi	-zu
	PL	-ri	-j	-w
Non-past	SG	-:za	-:zi	-:zu
	PL	-:ri	-j	-w

Table 36. Inflectional class for the made-up paradigm AVA

ERD		1	2	3
Past	SG	-da	-di	-du
	PL	-ri	-gi	-gu
Non-past	SG	-ga-za	-ga-zi	-ga-zu
	PL	-ga-ri	-gaj	-gaw

Table 37. Inflectional class for the made-up paradigm ERD

This type of change is well attested, and is recognized as one of the major ways by which multiple inflectional classes arise out of sound change (see in particular Wurzel 1987; see Chapter 4 for more detailed examples). The standard assumption is that sound change operates arbitrarily, and creates alternations that get morphologized.

3.1.2 Inflectional class from rule changes

Sometimes it does not take a real sound change for a sound alternation to create an inflectional class. Instead, stable sound patterns inside morphological paradigms associated with changes in the general phonological rules of the language can have similar effects. The following paradigms

(Table 38 and 39) show regular forms for a made-up language in which there is a regular rule voicing voiceless stops in between vowels. All the forms can be accounted for within the phonological rules of the language, with a regular alternation of phonetic realization of phonemes, such that /t/ > [d] /V_V, and remains [t] in all other contexts.

ABA	1	2	3
SG	aba-da	aba-di	aba-du
PL	aba-ri	aba-gi	aba-gu

Table 38. Made-up paradigm of the noun ABA

ART	1	2	3
SG	art-ta	art-ti	art-tu
PL	art-ri	art-ki	art-ku

Table 39. Made-up paradigm of the noun ART

If a rule change were to happen in that language, and the rule voicing voiceless stops between vowels ceased to be productive and part of the set of phonological rules of the language, but that for various reasons (potentially having to do with the fact that segments at affix boundaries tend to be shorter in duration), the forms inside the paradigm would not change, this could give rise to a distinction of inflectional class, if any noun not presenting the triggering context were to acquire the other set of markers, that is if these markers were reinterpreted as marking class features as much as inflectional values. In case of the alternation being still linked strictly to an alternation of phonological context, one could obviously argue that this would qualify as a case of construction specific phonology, or morphologically conditioned phonology (see Inkelas 2014:29-44).

3.1.3 Inflectional class from the loss of conditioning environment for alternations

The third possibility has to do with sound mutations triggered by a conditioning segment. These alternations tend to target the stem. Here is an example of a made-up language presenting such alternations. In this language, final stem vowels are stressed and are fronted when the following affix presents a front vowel. This is a simple process of assimilation (metaphony, quite similar to the real metaphony that happened in German or Italo-Romance). The following paradigms show the

original forms, followed by the new, alternating forms of that hypothetical language (Table 40):

	ABA (stage1)			ABA (stage2)
1	aba-tu		1	aba-tu
2	aba-ti		2	abɛ-ti
3	aba-rɛ		3	abɛ-rɛ

Table 40. Metaphony in a made-up paradigm

In stage 2 of the language, the alternation is still motivated. Now, if it happened in the language that final, unstressed vowels were centralized as schwa, the context triggering metaphony would be lost, but the stem alternation is still be retained as a sole marker of inflection, as shown in Table 41.

	ABA (stage3)
1	aba-tə
2	abɛ-tə
3	abɛ-rə

Table 41. Metaphony morphologized in a made-up paradigm

This type of change is basically what happens for the German class of plural of the type *Bruder ~ Brüder*, *Vater ~ Väter*, or what happened on stems for many Italo-Romance verb forms as well as nouns, giving rise to specific patterns of stem alternation. More will be said about metaphony in section 4.1.3.

It is perfectly possible to imagine metaphony being triggered by the stem vowel and giving rise to alternations in the form of affixes (most of the literature seems to describe this as the birth of vocalic harmony rather than metaphony), or applied to specific internal affixes and creating thematic markers. Thus, if two groups of denominal verbs were formed in a language by means of affixation with a different vowel, appearing in front of an affix causing metaphony, there would be an alternation ready to be morphologized: conditioned allomorphy is here reanalyzed as signaling class rather than as a purely phonological alternation.

3.1.4 Frequency conditioned reductive sound change

It is a well known fact that high token frequency items are more prone to phonological reduction than low frequency items (Bybee 2007). There is thus the possibility that a class of very high

frequency lexemes would reduce in similar ways to give rise to an inflectional class.

The case analysed by Dammel (2011) is complex, in that it involves such a class of phonologically reduced verbs (*Kurzverbklassen*), but it also involves some grammaticalization. In effect, the class of short verbs in the dialect of Bosco Gurin is by no means homogeneous inflectionally. Present forms realize person and number combinations through affixes that are formally similar for the whole class (though see the different patterns for 2PL in 'go' and 'stand'), but some of the verbs also present stem allomorphy patterns of different types (Dammel 2011:261; Table 42 shows only present indicative forms). In addition, the subjunctive forms are formed differently for those various verbs (Dammel 2011:261). One should note that those verbs belonged to different inflectional classes in Middle High German: they have thus converged to similar (though not identical) affixal realizations following phonological reduction.

Infinitive	sin	tya	gɒ:	ʃtɒ:	ha	ga	ge:	kse:
Meaning	'be'	'do'	'go'	'stand'	'have'	'give'	'take'	'see'
1SG	be	tyɑ	gɒ:	ʃtɒ:	hɑ	gep	ge:	kse:
2SG	beʃt	tyɑʃt	geiʃt	ʃteiʃt	hɛʃt	geʃt	ge:ʃt	ksiɑʃt
3SG	eʃt	tyɑt	geit	ʃteit	hɛt	get	ge:t	ksiɑt
1PL	siv	tiav	gɑ:v	ʃtɑ:v	hɛv	gɑ:v	ge:v	kse:v
2PL	sit	tiat	geŋgɑt	ʃtɛndat	het	gɑt	ge:t	kse:t
3PL	sen	tian	gɑ:n	ʃtɑ:n	hen	gɑ:n	ge:n	kse:n

Table 42. Short verbs in Bosco Gurin (Dammel 2011:261)

Dammel considers those verbs as a unitary class in that they all exhibit an innovative marking *-v* to mark first plural forms, an innovation that is not present in other classes. Dammel compares this innovative marking to the innovative marking arising in Gressoney, whereby subject pronouns are grammaticalized as new suffixes on verbs, giving rise to a potential new class (Dammel 2011:263). As such, she does not prove that reductive sound change in itself can give rise to a coherent inflectional class: the sound reduction seems to have given rise to a number of small membership classes, which happen to be the only ones innovating some affixal marking for the present via grammaticalization.

Frequency conditioned reductive sound change more often creates irregular lexemes which do

not pattern in a way found in any other paradigms in the language[22]. Such developments are very common, and may still be considered as instances of inflectional classes which only have one or few members. Thus in Gumbaynggir for examples (Pama-Nyungan, Australia, Eades 1979), there is only one major class of verbs, which all inflect in the same way, but there is also a number of very high frequency verbs which inflect in different ways (Eades 1979:301): 'go', 'take', 'hit', 'see', 'hear', 'say' and 'cry'. These irregular verbs might reflect earlier conjugations (Eades 1979:298), but one should note that most of them present a monosyllabic root unusual for verbs in that language, and may indicate some degree of phonological reduction. Similarly, in Watjarri (Pama-Nyungan, Australia, Douglas 1981), there are two main inflectional classes for verbs, plus seven irregular verbs with monosyllabic roots, which may indicate phonological reduction, as verb roots are generally disyllabic (Douglas 1981).

More will be said about the creation of inflectional classes from sound change in Chapter 4, with real-language examples.

3.2. Reanalysis

Writing about syntactic change, Harris and Campbell (1995) define reanalysis as "a mechanism which changes the underlying structure of a syntactic pattern and which does not involve any immediate or intrinsic modification of its surface manifestation" (Harris and Campbell 1995:61). Reanalysis also applies to morphology, where it can be understood as "a new way in which speakers understand the structure of a word by relating it to other words in a different, novel way" (Haspelmath 1995). This reanalysis of the internal structure of words (a change in their underlying structure) affects primarily phonological segments at affix boundaries and has to do with the problem of word segmentation. In effect, a segment or group of segments that were originally part of the stem come to be reanalysed as being part of an affix. The change does not affect the surface

22 Such irregular lexemes can still be considered as inflectional classes in themselves, but they do not serve to predict the inflection of other lexemes as would be the case for classes with at least two members. See Chapter 2.

structure of the form, only the way in which it is segmented by speakers. If phonological material belonging to the stem is ascribed to an inflectional affix, this means that different lexical items will now select different affixes. The change creates inflectional classes.

Haspelmath (1995) distinguishes three subtypes of morphological reanalysis: affix telescoping, conglutination, and secretion. Although most of his examples come from derivational morphology, it is possible to at least apply conglutination and secretion to inflectional morphology as well. These last two subtypes are also analysed by Haspelmath as cases of haplological overlap, whereby part of an affix (boundary) overlaps with either an inner affix (conglutination) or the root (secretion). Only those types of reanalysis involving haplological overlap are of interest in the case of inflectional class creation, and in particular secretion.

Secretion is the term used by Haspelmath for "the cases where a non-affixal part of a root is reanalyzed as part of an affix (Haspelmath 1995:8). It corresponds to what Dammel (2011) calls reanalysis. Let us consider the following made-up paradigm of nouns alternating for the feature number with only two values, singular and plural (Table 43). In such a paradigm, the realization of the feature-value plural is constant across lexemes.

Singular	Plural
alat	alat-es
ert	ert-es
jab	jab-es
actal	actal-es

Table 43. Reanalysis, stage 1

Very often, changes undergone by the base when in isolation can be seen as an example of morphotactic opacity, which according to Haspelmath (1995) is a major trigger for reanalysis, together with affix syllabicity. One can for example imagine that final consonants are lost, which triggers reanalysis, giving rise to paradigms in Table 44. As Haspelmath puts it, "in such cases where the base form is reduced phonologically, it is usually reinterpreted by speakers as the underlying stem form, and the affix which preserves the original stem consonant is reanalyzed as being part of the affix" (Haspelmath 1995:16). Such cases, as shown in Table 44, thus give rise to

allomorphy of the affixes, which implies that inflectional classes are born, which can then be extended to novel lexemes. In this case, because some of the earlier stem final consonants are identical, only three classes arise out of four lexemes.

Singular	Plural
ala	ala-te
er	er-te
ja	ja-be
acta	acta-le

Table 44. Reanalysis, stage 2

This case is similar to what happens in Maori, in Manam or in Toqabaqita (see Chapter 5).

Cases of conglutination can also in theory give rise to inflectional classes, although I have not found real convincing examples. One needs to posit a situation where two affixes occur together with a root, the inner affix being a derivational affix, and the outer affix an inflectional one. If conglutination occurs (the two affixes are not separable anymore), this gives rise to new allomorphs, and ultimately to a new inflectional class. Consider the made up forms in Table 45. These are forms of verbs for 1SG subject agreement, some with a derivational suffix, *-ma-*, say a suffix adding an inchoative aspectual distinction to its base (to begin to do something), and some without it:

Stem	Inflected form (1SG subject)
aba	aba-ti
aba-ma	aba-ma-ti
e-ma	e-ma-ti
car-ma	car-ma-ti
ela	ela-ti

Table 45. The prerequisite to conglutination in a made-up language

If the derivational suffix becomes semantically bleached, and if the inchoative value is not obviously present in the lexemes it derives, for example by presenting a more continuative meaning (to begin and continue doing something), there are high chances for conglutination to occur: the two suffixes coalesce into one, which can then be applied as an inflectional affix to novel lexemes without adding an inchoative meaning to its base (a new form *ela-mati* for example). One is now

left with two inflectional classes, one which realizes 1SG as *-ti*, the other which realizes it as *-mati*. Such conglutination is obviously helped in cases where the two affixes are becoming phonologically more fused, following phonological reduction or sound change, but such reduction is by no means necessary, contrary to semantic bleaching of the inner affix, which is a necessary component for inflectional classes to arise in such a way.

In Chapter 5, I analyse the creation of inflectional classes through reanalysis of the secretion type, examining examples from Maori, Toqabaqita, Manam, Romanian, and Nyawaygi.

3.3. Grammaticalization

Dammel (2011) explains that the creation of inflectional classes in Germanic languages through grammaticalization is not attested in the textual documentation of Germanic. It can still be reconstructed. Many studies look at the progressive morphologization of meaningful contrasts, and at the formal marking from analytical marking through full (functional) words to fully inflectional systems via various stages of cliticization and phonetic erosion (see Haspelmath 2011; Heine and Kuteva 2007). In order for grammaticalization to give rise to inflectional classes, one could assume that the formal change should develop to affixation. A number of studies show that grammaticalization can create inflectional classes in this way. This is in particular argued by Harvey (2008) who traces the origin of inflectional class markers in Guugu Yimidhirr from previous TAM markers. The fact that a marker that has meaning, like a TAM marker, becomes simply a marker of conjugation, is in fact an endpoint of grammaticalization: the marker is finally completely semantically bleached. What Dammel calls categorial reinterpretation (situations in which markers that are already present come to signify new distinctions by being interpreted in new ways) is in fact the last possible stage of any possible grammaticalization path, before the simple disappearance of the marker.

One can distinguish two possible scenarios for grammaticalization to give rise to inflectional

classes. One involves the grammaticalization developing to the creation of new affixes, a case that Haspelmath (1995) more properly calls agglutination. The other one is a simple process of grammaticalization giving rise to an auxiliary.

The first case of grammaticalization giving rise to inflectional class distinction involves the grammaticalization of periphrasis, which I analyse here from the example of auxiliaries. If a given language already presents inflection, and it concurrently develops auxiliation for a restricted class of verbs, then the distinction between verbs taking an auxiliary for their conjugation and verbs not needing it will be a distinction of inflectional class: two classes of lexemes that realize inflectional feature-values in different formal ways.

Agglutination can give rise to inflectional classes when there are cycles of grammaticalization occurring for the same features. This is particularly possible for the grammaticalization of subject (or object) agreement markers on verbs from originally free pronouns which become cliticized and bound to the verb. One can imagine a language going through two such cycles in succession, with the free pronouns being bound differing in form. If different lexemes grammaticalize different bound forms, this will give rise to inflectional class distinctions. The following made up language illustrates that. In example 1, there is a free pronoun and a verb, without agreement. In the next stage, the free pronoun has coalesced with the verb to give rise to an agreement marker which can now co-occur with another form of the free pronoun (Example 3).

(1) a baba

(2) a-baba

(3) b a-baba

In a next stage, the new free pronoun (*b* in Example 3) is also grammaticalized as a bound form with a different lexeme, to give rise to a new inflectional class, as shown in Example 4:

(4) b-aca

This cycle of similar grammaticalization for the same feature-value pairs has indeed arisen in some Papuan languages such as Skou (see Chapter 5).

Cases of inflectional classes coming into being via a process of grammaticalization are treated in Chapter 5. That chapter in particular distinguishes processes of auxiliation, such as those seen in Romance auxiliary selection, and the rise of an auxiliated conjugation in Basque, from cases of agglutination, as seen in the grammaticalization of subject and object agreement markers in the Papuan languages Skou and Arapesh.

3.4. Inflectional classes from etymological classes (or lexical strata)

While it might well seem extremely marginal, one can imagine getting inflectional classes from etymological strata in a language, in particular through the massive borrowing of inflected forms into a language. This would be generally possible for paradigms comprising a very small number of cells, typically nominal paradigms only associated with number values. For example, one could think of the high number of Latin words borrowed into English together with their plural forms as such a case. Thus new classes realising singular/plural pairs would be created in the language, the proof being the extension of the pattern to words that are not taken from Latin, but would find themselves assigned to a Latinate class of plural. Such words as *corpus* / *corpora*, *locus* / *loci*, *stratum* / *strata*, sometimes serve as templates for innovative plurals, for example the attested *octopus* / *octopi*. This shows that a new inflectional class is productive in the language, and that a plural in *-i* is as much a possibility for nouns presenting a singular in *–us* as the *–uses* plural. In such a case, it seems that a full paradigm has been borrowed from a foreign language, or rather, that the relevant cells of the paradigm (default cells for the feature-value pairs accessible in the language that borrows) have been borrowed, thus creating a potential paradigm. This might be the reason why such situations arise more easily if the borrowing language inflects according to fewer features than the source language, and if the borrowing language presents small paradigms. One could expect this to happen for languages where one or two features are marked with only two possible values (e.g. English, marking only number with values <sg, pl>, or Spanish marking gender and

number).

A specific case of inflectional classes due to language contact or borrowing comes from pidgins and creoles. These languages have often been characterized as either lacking inflectional morphology or at least as showing systematically less inflectional morphology than their lexifier (Roberts and Bresnan 2008). A number of cases seem to indicate that inherent inflection is better preserved from the lexifier to the pidgin or the creole variety, potentially because it is more closely associated with semantic values, not only with formal links of agreement betwen parts of the sentence or phrase (Roberts & Bresnan 2008:271-72). Roberts and Bresnan found that inherent inflection is retained twice as much as contextual inflection in the 29 genetically diverse pidgins of their sample. If this is the case, one could expect elements such as inflectional class to be less likely to be kept (or taken on) in pidgins, because they do not show a particular association with semantic values. That said, it is quite common for pidgins to lexicalize previous inflectional markers without retaining their associated semantics, giving rise to what Roberts and Bresnan call an 'empty word class marker' that is contrastive with other word classes, or sometimes fully integrated into the stem and no longer contrastive (2008:277-78). These lexified elements, though, seem only to mark the word category, and not to create subcategories of word classes which could give rise to inflectional class distinctions. There are only a few cases where inflectional class distinctions with multiple inflectional classes seem to be retained or later (re)innovated. Most of these have to do with nominal classification, where nouns exhibit different inflectional classes according to their gender assignment (see section 2.8.). In a number of pidgins, some class distinctions are kept and formally marked on the noun, although there are always fewer distinctions than in the lexifier language, and retention of inflectional classes is generally associated with a loss of the agreement system formerly marking grammatical gender relations. Thus in Fanalago and in Kenyan Pidgin Swahili, the system of noun classes is reduced from 15 classes to six, in association with a loss of agreement (Heine 1973:185-186); in addition, these classes are of the typical Bantu type, and are associated with the marking of singular and plural. Inflectional classes are thus essentially kept on nominals for the

marking of number, in a system with two values for the number feature (see Chapter 8 for a discussion of innovation of inflectional classes from gender systems). There seems to be a possible link with the acquisition of new classes from etymology, as seen above for English, which often targets nominals in their marking of number values. Thus in Broken Ojibwe, a gender distinction is only retained in the plural (3PL only, Roberts and Bresnan 2008: 279), which is an additional element in favour of a better retention of inflectional features for the expression of (nominal) number. It is possible that inflectional classes are retained or created on noun classes because these classes are more closely associated with basic semantic properties (such as being human, or edible).

In at least one case, an inflectional class distinction is kept for verbs, and is mainly used for distinguishing verbs taken from two different strata of the lexicon of the creole in question (see Luís 2011). This seems to be less common than the retention or subsequent acquisition of inflectional classes for nominals, probably because verbal paradigms tend to be larger than nominal paradigms. Luís (2011) presents the case of two Portuguese lexifier creoles, Darman and Korlai Portuguese Creoles. There again, paradigms are restricted to a small number of distinctions, yelding a total of four cells per paradigm. But the two languages still present inflectional classes: three are inherited from Portuguese, and one is a novel development and is exclusively used to accommodate loan-words. This is thus a case of lexical strata involving a new inflectional class distinction.

In cases where an inflectional marker is completely lexicalized and fused with the stem without bearing any more meaning, one could imagine that such distinctions could give rise to inflectional classes, in particular through favouring the emergence of a thematic element. This is mainly possible if not all lexemes of a given word class bear that marker. For example, in Gulf Pidgin Arabic, around half of the verbs present a prefix *y(V)-* normally associated with 3SG subject in Arabic. If in a future development the language were to develop inflectional markers by means of prefixes (say, bound subject person/number markers, as the language is SV), a distinction would arise between stems presenting stem initial yod, and the other stems, if only for morphonological reasons, and would potentially give rise to inflectional class distinctions.

Finally, some languages have a systematic class assignment for loanwords. While this does not properly create a class, it can certainly reinforce specific classes. Thus, in Biak, all loanwords are adapted into verbs by using the verbalising element *be-* inflected like a second class verb[23]. This creates a specific class of verbs all having the same prefix (a derivational class), and all belonging to the same inflectional class (class 2, infixal). Although indigenous words also take this verbalising prefix, this certainly creates a substantial correlation between loanwords, a derivational class and inflectional class, even more with the extensive borrowing due to contact and massive bilingualism with Indonesian and varieties of Malay. Correlations between derivational classes and loanwords are commonplace, as is the case in English where there is a series of older suffixes and a series of more recent, Latinate suffuxes borrowed from French, whose use depends on the stress patterns of the base.

3.5. Heteroclisis

Heteroclisis is the morphological phenomenon by which the paradigm of a given lexeme appears to follow one inflectional class for a subset of its cells, and another paradigm for another subset of its cells (Stump 2006; Kaye 2015). As Kaye (2015) states, "what marks out the paradigm of a heteroclite lexeme, or group of lexemes, is that it does not pattern exclusively with any single one of the inflectional classes which can otherwise be identified for the language – but instead contains some cells which show morphological behaviour proper to one of these classes, alongside other cells which show morphological behaviour only proper to others. In an instance of heteroclisis, a single paradigm thus seems to be 'split' between two or more established inflectional classes" (Kaye 2015:2). If one follows this definition, a number of inflectional classes traditionally recognized as such have to be redefined as heteroclitic, as is the case with the Latin third declension of nouns.

23 The verbalizer *be-* is originally a verb belonging to the second class.

The issue of heteroclisis in the context of the origin of inflectional classes is tightly bound to the type of conception one has of an inflectional class. If one follows Stump's (2016) hypothesis that inflectional classes are classes of *stems* inflecting in a particular way, not classes of lexemes as usually thought (e.g. Aronoff 1994), heteroclisis is easily formalized. A lexeme is made up of two stems, potentially homophonous, which each select a different inflectional class. In such a conception, heteroclisis does not create a novel inflectional class in the language, it simply parallels the inflectional properties of the language in general. But, as stated in Chapter 2, it often leads to positing homophonous stems which are only distinguished by their inflectional behaviour. Thus in the case of the Latin noun *domus* (Table 46), one must specify at least two homophonous stems: one that inflects according to the model *servus* (second declension), one that inflects according to the model *manus* (fourth declension). In addition, one should note that it also leads to positing such stems for the other lexemes, as there would be a discrepancy in the cells that select the *servus* inflectional class for *servus* and *manus* (the cells following the second declension in *domus* should also be the only cells following it in *servus*; if that was not the case, the two inflectional classes would be different in that their content paradigm would look different; see Section 2.3. on that issue). This leads to positing a morphomic pattern of stem selection rather than stem alternation which does not correspond to any actual alternation in form.

	'slave'		'house'		'hand'	
	SG	PL	SG	PL	SG	PL
NOM	servus	servī	domus	domūs	manus	manūs
ACC	servum	servōs	domum	domōs	manum	manūs
GEN	servī	servōrum	domūs	domōrum	manūs	manuum
DAT	servō	servīs	domuī	domibus	manuī	manibus
ABL	servō	servīs	domō	domibus	manū	manibus

Table 46. Heteroclite inflection of Latin domus *'house' (Kaye 2015:2)*

On the contrary, if one considers that an inflectional class is a class of lexemes which realize inflectional values in the same way, as I do in this thesis, a heteroclite lexeme presents its own inflectional class which can be said to inherit from two different inflectional classes in an inheritance hierarchy model. In such a model, heteroclisis is recognized as a specific phenomenon

in that it is the only instance where a given node inherits from two higher nodes rather than one. But in terms of inflectional class, it has to mean that a heteroclite lexeme is of a different inflectional class than its two higher nodes are. If one recognizes a heteroclite lexeme or set of lexemes inflecting in the same way as a separate inflectional class, then inflectional classes can arise from heteroclisis.

How does heteroclisis arise? Maiden (2009) shows that heteroclisis arises because of inflectional class ambiguity in subdomains of the paradigm, in the Romanian example Maiden takes due to sound changes. This ambiguity leads speakers to reanalyse part of the paradigm as belonging to a different inflectional class. Maiden (2009) has shown that heteroclisis tends to develop on the lines of existing patterns of stem allomorphy. This would imply, as in Stump's (2006, 2016) theory, that a different stem selects a different inflectional class in instances of heteroclisis. In most cases of heteroclisis treated in diachrony available in the literature (see Kaye 2015:111-127), what happens is that a given lexeme following a given inflectional class undergoes an incursion of new forms following another inflectional class for parts of its paradigm. This incursion follows the lines of established partitions of the paradigm in the language. In some cases, heteroclisis can be shown to be the result of a progressive change of inflectional class: a given lexeme moves from one class to another in a progressive way (Esher 2012; see also Bach & Esher 2015).

3.6. Deponency

Another example of a morphological phenomenon that can give rise to inflectional classes is deponency. Deponency is a morphological mismatch between the expected form of a paradigm for a given set of feature values, and its actual realization, which appears to correspond to the usual realization of different values. The main exemplar of deponency concerns Latin verbs. Baerman (2007) defines deponency in Latin as follows: "Deponency is a mismatch between form and function. Given that there is a formal morphological opposition between active and passive that is

the normal realization of the corresponding functional opposition, deponents are a lexically-specified set of verbs whose passive forms function as actives. The normal function is no longer available." (Baerman 2007:1). Thus, Latin deponent verbs formally exhibit passive morphology, corresponding to the marking usually present for passive voice, but this marking in fact realizes active voice, creating a mismatch between what is marked and the way it is marked. Table 47 presents the present indicative paradigm of three verbs: *amo* 'love', an active verb; *amor* 'to be loved', its passive counterpart; and *loquor* 'speak', a deponent verb. As is apparent in the table, *loquor* exhibits passive morphology, although it is an active verb.

	'love'	'be loved'	'speak'
1SG	amo	amor	for
2SG	amas	amaris	faris
3SG	amat	amatur	fatur
1PL	amamus	amamur	famur
2PL	amatis	amamini	famini
3PL	amant	amantur	fantur

Table 47. Active, passive and deponent verbs in Latin

There are deponent verbs belonging to the four main conjugations of Latin: first conjugation *cōnor* 'try, attempt', second conjugation *polliceor* 'promise', third conjugation *loquor* 'speak, talk', fourth conjugation *orior* 'rise, arise'. One should note that deponent verbs innovated a few (non-finite) forms regularly formed like normal active verbs: the present and future active participle, and the future active infinitive. The morphological mismatch is thus only present for a subset of their paradigm.

Deponent verbs have not generally been considered as new inflectional classes, probably because of their exceptionality: they are already singled out in the description as being exceptional. In addition, they take most of their forms from existing inflectional classes. But if one follows a definition of inflectional classes in which both the content paradigm (list of feature value bundles) and the form paradigm (list of realizations) have to be identical for two lexemes to belong to the same class, then one has to recognize deponent verbs as creating new inflectional classes: they share a form paradigm with passive verbs, but they share their content paradigm with active verbs, except

for the fact that they do not have a passive (thus, their content paradigm is defective). Deponency would thus be another possible case of inflectional class creation.

3.7. Inflectional classes from alienability distinctions

Alienability distinctions are a type of nominal classification restricted to possessive paradigms that typically oppose two classes. One class is said to be a class of inalienable items, referents whose possession is seen as more permanent, and the other a class of alienable items whose possession is seen as more transitory. A large number of languages present alienability distinctions that are sometimes argued to be based purely on the semantics of the nominal referent, in particular in Oceanic languages[24]. This type of classification is found on all continents. Inalienable items typically include body parts and kin terms.

In addition to languages where the distinction between alienables and inalienables is purely based on the semantics of the referent, a large group of languages presents a situation where the membership of the closed inalienable class is somewhat restricted lexically: although most members are body parts or kin terms, not all body part terms and kin terms belong to this class. In such cases, the semantics of the nominal referent are not enough to characterize membership of the inalienable class: the class membership has to be specified lexically. This inalienable class thus forms a lexical class whose main characteristic is to realize the marking of possession in a different way. Lexical classes whose members realize some inflectional feature in the same form are inflectional classes. One can thus consider such cases as inflectional classes which arose from an original system of alienability distinction, which amounts to saying that they arise from a distinction in lexical semantics coupled with some distinction in the marking of possession.

In some languages, class differentiation has gone further, and the systems present a larger

[24] It seems that such purely semantics based systems are in fact very hard to find, if at all present (Greville Corbett p. c.).

number of inflectional classes. Typically, such systems oppose a number of small classes with restricted membership, mostly reserved to lexical items which would be part of the inalienable class, to a single open class with large membership which corresponds to items that would traditionally be treated as alienable.

In Chapter 7, I examine a number of languages where such developments have arisen. In particular, a number of languages of the South Halmahera-West New Guinea family of Austronesian languages present such developments, but they are also found in other Austronesian languages, as well as in non-Austronesian languages of Papua (Anêm, Amele, and Salt-Yui), in Worroran languages of Australia, and in the Pomoan and Cochimi-Yuman languages of Northern America.

3.8. Inflectional classes from agreement class (in particular gender)

It is in theory possible to acquire inflectional classes from gender systems in two different ways. Gender agreement systems include both a controller – a noun in this case – and targets that show a systematic covariance in form with the controller: for any given controller, a similar target will vary in the same way. Obviously, targets can covary according to their own lexical class and inflectional class. The interesting point here is that these two pathways to inflectional classes are clearly separate, and show a neat divide between inherent and contextual markers.

On one hand, it is possible to have inflectional classes on the targets of gender agreement. A classic example of that situation is found in Latin, where adjectives agreeing with a controller noun in gender and number systematically covary with the noun according to their own inflectional class: a given adjective will show a consistent form in a similar syntactic function when associated with nouns of the same gender. But another adjective taken from a different class, although it will still show a similarly consistent behaviour, will show different exponents.

On the other hand, some languages mark gender on the controller in addition to showing

agreement on the target. This is a situation typical of Niger-Congo languages. Traditionally, this marking has been said to involve noun class, and in most languages it correlates in a partial way with agreement classes, although none of the languages examined presented a consistent match between noun class and agreement class markers. But the marking of controllers is not strictly speaking gender marking, nor does it have to correlate systematically with the agreement classes defining genders. In fact, it would be far more precise to distinguish them, and to refer to the controller marking as inflectional class, as it is also correlated with the inflectional marking of number. Cases where these classes strictly correlate with gender classes are simply another illustration that an inflectional class can correlate and interact with other components of the grammar: in this case, inflectional classes interact with gender systems understood strictly as agreement classes. Inflectional classes and gender systems simply tend to align for maximising the reliability of formal inferences, meaning that the knowledge of the inflectional class of a noun will reliably correlate with the gender for which it will control agreement.

Evidence for such a claim comes both from languages where the correlation between inflectional class and genders is not perfect, and from languages that have lost gender agreement in part or totally.

Still, in languages where there is a strong correlation between agreement markers and gender/class markers on nouns, both using the same forms, one can argue that the inflectional class systems of nouns arose from the gender system marking. This is in particular the case for languages showing alliterative concord, such as Niger-Congo languages or the Arapesh languages of Papua New Guinea. In those languages, a system of inflectional class can be shown to arise through a combination of gender marking on the noun with sometimes identical realizations to the marking on the agreement targets, and the fact that nouns inflect for number, which covaries with the gender marking[25].

A large number of languages in West Africa show a reduced agreement system, where the

[25] Nonetheless, in synchronic terms, no language has yet been found where such coreelation holds for all forms. In Arapesh for instance, only singular forms show alliteration, not plural forms.

correlation between inflectional class on nouns and the number of genders shown through agreement is now far from perfect (ex. Gur languages, see Chapter 8). In yet other languages, the agreement system has disappeared, leaving only inflectional class marked on nouns. The loss of gender agreement is documented for Niger-Congo languages in the Gur subgroup: Dagaare (Bodomo 1997; also see Grimm 2010), Lobi (Greenberg 1978:52), Chakali (Brindle 2009).[26] More will be said about such systems in Chapter 8, where I analyse the development of inflectional class from gender systems in the Gur languages, and the correlation between gender and inflectional class on nouns in some Papuan languages (see Aronoff 1994).

3.9. Impossible classes: No-Blur Principle and Low Conditional Entropy Hypothesis

After discussing the possible origins of inflectional classes, I now turn to the related issue of the principles which may constrain such classes, and outline a system of impossible classes. Allomorphy has been treated as limited by a number of possible constraints, of which the No Blur Principle (Carstairs-McCarthy 1994) is the main exemplar. It is has more recently been called Vocabular Clarity (see Carstairs-McCarthy 2010; Enger 2013). Such a Principle makes predictions concerning the range of possible and impossible inflectional class systems. The Principle has been subject to a number of criticisms in recent literature (Müller 2007; Halle & Marantz 2008; Stump & Finkel 2013; Ackerman & Malouf 2013, 2015). This section explains the No Blur Principle, and some of the criticisms it has received, and gives a counterexample to it from a Pama-Nyungan language of Australia, Mpakwithi (Crowley 1981). I show in particular that very simple inflectional systems can contravene the No Blur Principle, by analysing the inflectional class system of Mpakwithi.

26 Interestingly, both Dagaare and Chakali have developed a new, non-canonical gender system, targetting exclusively 3PL forms, and based on animacy.

3.9.1 No Blur Principle

The No Blur Principle is a principle constraining the distribution of inflectional affixes in an inflectional class system, first introduced by Carstairs-McCarthy (1994; see also Cameron-Faulkner & Carstairs-McCarthy 2000) as the corollary for inflectional morphology of the more general Principle of Contrast proposed by Clark (1987). Clark's (1987) Principle of Contrast states that "any two forms contrast in meaning". Carstairs-McCarthy states that for inflectional morphology, such meaning must be understood as containing information about inflectional class membership, so that either an affix has precise information about the inflectional class the inflected lexeme belongs to, or it is a default realization for the cell realized. The derived principle, termed the No Blur Principle, thus states that: "Within any set of competing inflectional affixal realizations for the same paradigmatic cell, no more than one can fail to identify inflectional class unambiguously" (Carstairs-McCarthy 1994:742).[27] That is, "within any set of competing classes, each affixal realization for every cell must be either a) peculiar to one class or b) the only realization for that cell which is shared by more than one class" (Carstairs-McCarthy 1994:741). The No Blur Principle thus predicts that a system in which two or more affixes happen to be used for the same cell by two or more inflectional classes is impossible, or at the very least, difficult to acquire (Carstairs-McCarthy 1994:743-744).

Carstairs-McCarthy (1994) gives evidence from a range of languages in support of the No Blur Principle, including Icelandic, German, English, Archi, Andi, Georgian, and Latin. He does not however claim that the principle is inviolable, only that such systems that may contravene it are at best difficult to acquire: thus most of the counterexamples must include high frequency items (as is the case in his English verb example). The No Blur Principle is different from a previous principle proposed by Carstairs-McCarthy, the Paradigm Economy Principle, which states that "at least for one cell, all the realizations should be class-identifiers" (Carstairs-McCarthy 1994:754). The two principles are obviously related but they differ in some major ways. In particular, the Paradigm

27 Note that Carstairs-McCarthy (1994) restricts the working of the principle to affixal realizations.

Economy Principle states that at least one cell should comprise only class identifiers, and no default marking, while the No Blur Principle does not go that far: all cells can be marked by defaults, as long as there is only one such default which is not a specific class identifier.

3.9.2 Exceptions to the No Blur Principle

Some of the counterexamples to the No Blur Principle given in the literature show extremely complex inflectional systems (in particular Nuer, Baerman 2012, or Fur, Stump & Finkel 2013). In the next subsection I show that even a rather simple system like that of Mpakwithi can contravene the No Blur Principle.

Nuer (Western Nilotic, South Sudan; Baerman 2012) has up to 25 inflectional classes for nominals, although it only presents three different affixal realizations for the various case and number combinations, *-kä*, *-ä*, and *-ni*. What creates the multiple inflectional classes is the various distributions of these affixes, which do not denote any specific bundle of features. That paradigms are blurred can be show from the four lexemes given as examples in Table 48.

	'stone'	'umbilical cord'	'peace'	'sky'
NOM SG	döl	ca̱a̱r	mal	puäär
GEN SG	döl-kä	ca̱a̱r-ä	mal-ä	puäär-kä
LOC SG	döl-kä	ca̱a̱r-ä	mal-kä	puäär-ä

Table 48. Blurred inflection in Nuer (Baerman 2012:469)

In Table 48, the suffixes *-ä* and *-kä* both can have the status of default for the genitive singular, as well as for the locative singular cells: both suffixes are present in more than one inflectional class. This is thus a clear example of blurring. Baerman (2012:471) also gives data on Latvian nouns which contravene the No Blur Principle.

For Ackerman & Malouf (2015), the No Blur Principle is just one of the possible arrangements that obtains when languages conform to their Low Conditional Entropy Conjecture. They refer to evidence that there are languages which do not conform to the No Blur Principle (Baerman 2012; Stump & Finkel 2013), and present a developed computer simulation of the evolution of the inflectional morphology of a selection of artificial languages. The Low Conditional Entropy

Conjecture is designed as a way to give a measure to the Paradigm Cell Filling Problem: given the knowledge of a word form, how reliable is the inference the speaker makes to produce another word form in the same paradigm? Conditional entropy is a measure of the reliability or of the uncertainty of such inferences, averaged out for the whole paradigm. The lower the entropy, the more reliable is the inference being made between forms in two different cells; the higher the entropy, the less reliable that inference. Entropy measures thus range from 0 (absolutely reliable, systematic inference) potentially to the infinite. The Low Conditional Entropy Conjecture was introduced in Ackerman & Malouf (2013), where the authors computed the mean conditional entropy for exemplar paradigms in a series of genetically unrelated languages, which yielded the result that most languages show a conditional entropy ranging from 0 to 0.75, i.e. that languages tend to have a low conditional entropy in the structure of their paradigms. Ackerman & Malouf (2015:6) indicate that a language conforming to the No Blur Principle will by definition have a low entropy, but that conversely one can conceive languages that infringe the No Blur Principle but conform to the Low Conditional Entropy Conjecture. Nuer is such a language (Ackerman & Malouf 2013). The computational modelling with iterative learning in Ackerman & Malouf (2015) yields such results: a significant number of the produced evolutions show blurring, but still conform to the Low Conditional Entropy Conjecture.[28]

3.9.3 The Mpakwithi dialect of Anguthimri (Pama-Nyungan, Australia)

Mpakwithi (Crowley 1981) is the only documented dialect of Anguthimri, a Pama-Nyungan language once spoken in Northern Queensland, Australia. Its nouns inflect for case only. The language presents split ergativity, whereby nouns inflect according to an ergative-absolutive alignment, butpronouns inflect on a nominative-accusative basis, which is common in Australian languages. It also shows vowel harmony, which will be relevant for the form of suffixes seen in nominal inflection. Once vowel harmony is factored out, most affixes only exhibit one form; only

28 The Low Conditional Entropy Conjecture does not set an upper bound for the entropy displayed by actual languages, but it simply states that it should be low.

the ergative and the genitive show allomorphy and are distinguished for the exponence of a number of inflectional classes: all other cases present default forms throughout all classes. The instrumental is always syncretic with the ergative.

The ergative / instrumental shows three possible allomorphs: *-gV*, *-rV*, and *-ɟ̱V* (where V stands for any vowel corresponding to the vowel harmony of that word). There does not seem to be any phonological conditioning for the choice of those allomorphs (Crowley 1981). In addition there are two subclasses of irregulars taking different allomorphs in the ergative, one where the last *-i* or *-e* vowel of the word is replaced by *-a*, which changes the vowel harmony; the other with a replacive morph: words ending in *-ɣi* being replaced by *-ɾ̠i* (e.g. absolutive *ni:ɣi*, ergative *ni:ɾ̠i*).

In the genitive, three allomorphs are present, *-mɾ̠a*, *-namɾ̠a*, and *-ɣamɾ̠a*. There is no phonological conditioning for the choice of those allomorphs, except that *-mɾ̠a* is restricted to monosyllabic roots. Monosyllabic roots being constrained by phonology, they need not concern us here[29].

For regular declension, we are thus left with three main allomorphs for the ergative and two for the genitive, that is a maximum number of six possible classes. Anguthimri comes close to exhibiting the maximal number of possible classes, in that five inflectional classes can be reconstructed from the data given by Crowley (1981)[30] as shown in Table 49, plus three irregular classes (Table 50). The tables only give the forms relevant for inflectional class definition, absolutive, ergative / instrumental, and genitive.

	1A	2A	2B	3A	3B
	'alligator'	'wild dog'	'tame dog'	'mosquito'	'wasp'
Absolutive	kyabara	ɖ̱walaɖ̱i	ʔwa	ŋu:lu	ɖ̱e:ni
Ergative	kyabara-ga	ɖ̱walaɖ̱i-ri	ʔwa-ra	ŋu:lu-ʈ̱u	ɖ̱e:ni-ʈ̱i
Genitive	kyabara-ɣamɾ̠a	ɖ̱walaɖ̱i-ɣamɾ̠a	ʔwa-namɾ̠a	ŋu:lu-ɣamɾ̠a	ɖ̱e:ni-namɾ̠a

Table 49. Regular inflectional classes in Mpakwithi (reconstructed after Crowley 1981)

29 Because the choice of inflection is constrained by phonology, such lexemes are not relevant for the purpose of the No Blur Principle, as part of the 'meaning' of the affix is chosen on non arbitrary grounds. Note that they still fall into two different classes for the marking of the ergative case.

30 Crowley (1981) does not give full tables of inflection for nouns. The data presented here is reconstructed from his rules about how forms are constructed, and from the indications of class for ergative and genitive given in the small lexicon accompanying his grammar.

	Irregular 1	Monosyllabic 1	Monosyllabic 2
	'boy'	'hand'	'foot'
Absolutive	niːɣi	ʔa	kwe
Ergative	niːɽi	ʔa-ga	kwa-ra
Genitive	niːɣi-ɣamɽa	ʔa-mɽa	kwe-mɽa

Table 50. Irregular inflection in Mpakwithi (reconstructed after Crowley 1981)

The lexicon given by Crowley (1981) only gives full inflectional information for a few lexemes. It is fairly possible that the unattested possible class existed but was not recorded. Table 51 gives the number of lexemes for which full inflectional information is known and their repartition by inflectional class (out of a total of 350 nouns recorded):

Class	Lexemes
1A	2
1B	0
2A	3
2B	2
3A	6
3B	5
TOTAL	18

Table 51. Repartition of classes in Mpakwithi

Crowley (1981) claims that it is not possible to find any phonological conditioning for the distribution of allomorphs into classes. These classes, although relatively few and simple, show extensive blurring of affixes. Two of the possible affixes for the ergative case compete for the status of default: *-rV* and *-ɻ̈V* are both present in two inflectional classes. Similarly, of only two affixes possible for genitive case, none signals its inflectional class unambiguously: *-ɣamɽa* is present in three classes, while *-namɽa* is present in two different classes. The inflectional system of Mpakwithi is thus a clear example of a language that infringes the No Blur Principle.

3.9.4 Impossible classes

The No Blur Principle is not sufficient to determine a range of impossible inflectional classes in language, as a number of languages infringe it, although their inflectional system can be quite simple, as it is the case in Mpakwithi. The Low Conditional Entropy Conjecture (Ackerman & Malouf 2015) can be used instead as a limitation on the possible systems that can arise. Specifically,

it prevents systems of classes with high conditional entropy to be maintained over time: such systems will always tend to be regularized to show a lower conditional entropy, as shown by computational models used in Ackerman & Malouf (2015). Languages following the No Blur Principle will also have a low conditional entropy; there is thus the possibility that a large number of languages still follow that principle, although it need not be so. In any case, the No Blur Principle cannot serve as a diagnostic for what would be impossible classes, for systems of inflectional classes showing blurring of affixes do exist.

3.10. Conclusion

This overview chapter shows that inflectional classes can originate in a large number of phenomena. In addition to those phenomena already referred to in the literature, including sound change, grammaticalization, and reanalysis, one can adduce further possible scenarios for the creation of inflectional classes. They can originate in lexical strata in the case of language contact. Because of the definition of inflectional class adopted in this work, one also has to consider that the two morphological phenomena of heteroclisis and deponency also give rise to new inflectional class distinctions. Finally, a system of inflectional classes can also originate in an already fully fledged system, in the form of alienability distinctions and gender systems. These phenomena are new additions to the diachronic typology of the origins of inflectional classes.

The following chapters develop some of these sources of inflectional classes, and expand the range of languages for which they are attested beyond Germanic or more generally Indo-European languages. The three main phenomena presented in the literature, sound change, grammaticalization, and reanalysis, are shown to hold for a wider range of languages as sources of inflectional classes. I then further develop the possibility for inflectional classes to originate in already existing systems of gender or of alienability distinctions.

4. Sound change as a source of allomorphy and the rise of inflectional classes

The main evolutionary scenario for the emergence of inflectional classes in language proposed by Carstairs-McCarthy (2010:102-109) has to do with sound change. For him, inflectional classes arise through the application of sound changes in paradigms, which give rise to allomorphy. He gives persuasive made up examples which show that even regular sound changes can give rise to systems that present non conditioned allomorphy (Carstairs-McCarthy 2010:103-104). This chapter treats the creation of new inflectional classes through sound change, which is the main source generally recognized for inflectional classes (see also Dammel 2011).

The scenario put forth by Carstairs-McCarthy (2010) is relatively close to what will be shown in section 4.3. for South Halmahera-West New Guinea languages. In each case, it has to do with the morphologization of a marker: in Carstairs-McCarthy's example the morphologization of a tense distinction, facilitated by some sound changes; in the South Halmahera-West New Guinea languages, the morphologization of pronominal agreement in verbs, through the cliticization and subsequent affixation of pronouns to the verb, followed again by some sound adaptation. In both cases, the result is the creation of inflectional classes. But the scenario in Carstairs-McCarthy (2010) is rather peculiar, in that it includes both a series of sound changes and some changes in the morphology: it presupposes a change in the segmentation of forms by the speakers. This is shown in the example he choses of the Maori passive conjugation. Maori displays, in its passive forms of verbs, what looks like an inflectional suffix *-ia*, preceded by various consonants which do not appear in the active form, as shown in Table 52 (transcription shown in IPA):

	Active	Passive
'drink'	inu	inu-mia
'catch'	hopu	hopu-kia
'embrace'	aɸi	aɸi-tia
'carry'	mau	mau-ria
'point out'	tohu	tohu-ŋia

Table 52.Active and passive verb forms in Maori (after Carstairs-McCarthy 2010:105)

The system of Maori verbs thus looks as if it shows a number of inflectional classes.[31] An alternative analysis, that some prefer, is to recognize that there is an identical marker *-ia* marking passive, and that the consonant is thus a thematic element only marking inflectional class, with a different segmentation, not *inu-mia* but *inu-m-ia*. In any case, the present situation of these consonants comes from sound change and a resegmentation by speakers. Active forms used to present a final consonant in the stem, which was deleted by sound change. That initial situation is shown in Table 53:

	Active	Passive
'drink'	inum	inum-ia
'catch'	hopuk	hopuk-ia
'embrace'	aɸit	aɸit-ia
'carry'	maur	maur-ia
'point out'	tohuŋ	tohuŋ-ia

Table 53. Reconstructed earlier system in Maori

When the final consonants were lost, their equivalent in the passive forms were not, protected by the passive suffix, that is by their phonologically non-final position. Speakers reinterpreted the last consonant of the stem as part of the desinential material, thus creating a large number of inflectional classes through a single sound change. This example thus involves reanalysis as much as sound change. I show in section 4.3. that a class creation happened when subject pronouns were morphologized as agreement markers in verbs in South Halmehare-West New Guinea languages. The phonetic adaptation to the shape of the root created some allomorphy. Further sound change created more classes.

There is thus a consensus in the literature that the main way in which inflectional classes are created in language is through sound change. Sound change is implicated in a majority of the new class creations in Germanic exemplified by Dammel (2011), although there are also other types. But one can imagine and find other types of sound changes than those exemplified by Carstairs-McCarthy (2010), which is more properly treated as a case of reanalysis (see Chapter 6) or at least

31 There has been some debate as to whether the Maori facts should be treated in the phonology or in the morphology (see Blevins 1994). However these facts are treated, they pose a problem for the analysis. A morphological analysis in terms of inflectional classes is here preferred.

imagine other types of phonologically triggered allomorphy (see Section 3.1.). In particular, one should distinguish the type of sound adaptation that occurs when affixes morphologize, which can sometimes create multiple classes (Section 4.1. and 4.3.), from what happens when sound change affects a paradigm that already exists and presents one or more classes (section 4.1.). In this latter case, classes are always created from existing classes, in effect creating subclasses which can be formalized in an inheritance hierarchy model.

4.1. Sound change and conditioned allomorphy

Sound change can create alternations which are recruited by the morphology to mark some distinctions. The morphology being expressed by strings of sounds (all morphology inherently presupposes phonological substance, concatenative or not), this is hardly surprising. As Koefoed & van Marle put it, "the morphologization of originally morphonological phenomena is not particularly rare. Morphological phenomena are concomitant i.e. they are associated with a particular morphological process. If the overt marker of that process (...) disappears, the concomitant phenomenon may take over the function of overt morphological marker" (Koefoed & van Marle 2000). This is for example what happened for some plural marking in German of the type *Bruder ~ Brüder* (section 4.1.3.), where a number of nouns also acquired a metaphonic alternation that was not etymologically present through analogy (old German *boum, bouma* 'tree', but modern German alternation *Baum, Bäume* 'tree'). While German extended the pattern, Dutch drastically levelled it, leaving only the exceptional alternation *stad, steden* 'city'.

But morphologization also applies to non phonological distinctions, in the cline of grammaticalization which goes from juxtaposed elements to affixes via a process of cliticization (Lehmann 1995, section 4.1.1.). Such gradual changes involve some degree of phonological adaptation at morph boundaries, which create morphologically conditioned sound changes. Such changes have consequences for the way paradigms will be organized.

4.1.1 Grammaticalization and boundary sound adaptation

The first situation in which some sound changes create inflectional classes is in the case of grammaticalization / morphologization. The grammaticalization path from a full lexical element to an affix through cliticization is well known, and may involve any or all of decategorialization, semantic bleaching, and phonetic erosion (see Lehmann 1995). A good example of such developments is the rise of synthetic future and conditional from an original Latin periphrasis in the Romance languages (Esher 2012), as shown by French: *cantare habeo* > *chanterai* (future); *cantare habebam* > *chanterais* (conditional).

When a new element starts to become a clitic and then an affix, there may be some sound adaptation going on at the morph boundary depending on the phonological properties of the stem chosen. There is a phonetic adaptation in concatenation which will particularly target stems ending or starting in a vowel as opposed to those ending or starting in a consonant (depending on the type of affixation, either prefixation or suffixation). This will give rise to conditioned allomorphy, with one form of the affix being attached to stems starting in a vowel and another form to those starting in a consonant.

A morph boundary is thus a preferred place for sound change to occur. This is shown both by phonetic evidence about the timing of segments at an affix boundary (shorter than usual) and from the occurrence of morphologically conditioned sound change.

Sound change must occur across an affix boundary to create inflectional class: a sound change occurring within the affix will not create any distinction between lexemes. A sound change occurring within the stem will only create an inflectional class if it is correlated with the presence of a specific affix (or type of segment within the affix), as is the case with metaphony (see 4.1.3). Any other type of sound change within the stem will only affect the class of stems that contain a specific segment, and will not correlate with inflectional distinctions.

4.1.2 Sound change inside an existing inflectional paradigm

This section examines cases of sound change giving rise to new inflectional classes when such sound changes occur at affix boundaries. I examine two such cases in the Romance languages, one in Francoprovençal, the other in Romanian dialects. Most Romance languages show either conservation of the four inflectional classes of Latin verbs, characterized by a specific thematic vowel, or a reduction thereof, at least in some parts of the paradigms.

4.1.2.1. Francoprovençal

Francoprovençal dialects (Romance, Indo-European; Kristol 2016) are spoken along the border of France, Switzerland and Italy, mostly in the Alps. In all Francoprovençal varieties, the Latin first conjugation (verbs having an infinitive in *-are*) has yielded two inflectional classes because of phonological change. In particular, verbs whose root ended in a final palatal in early Romance form a separate subclass of first conjugation verbs, differentiated in the imperfect indicative and infinitive in all varieties (Kristol 2016; Bjerrome 1957:86). The dialect of Bagnes (Bjerrome 1957:87-89; Table 54) differentiates those classes in the imperfect indicative, imperfect subjunctive, conditional, second person plural present indicative, and infinitive.

The development of those two inflectional classes out of the single first class of Latin verbs has to do with a differential treatment of Latin stressed /a/ in open syllables. The first conjugation of Latin is characterized by the appearance, in many parts of its paradigm, of a thematic vowel /a/, which happened to bear stress in the infinitive and the imperfect in particular. Latin stressed /a/ was fronted after palatal consonants as early as the sixth or seventh century CE (Kristol 2016:353).[32] A similar development affected final unstressed /a/ (Kristol 2016:353). This development affected first conjugation verbs, and as a result it split that conjugation into two separate subclasses (Table 54).

32 Palatal consonants did not exist in Latin and were a development of proto Gallo-Romance. Velar consonants followed by /a/ in particular palatalized, which explains why *mandu'kare 'eat' evolved to *man'ʤar which then underwent the vocalic change.

PRESENT	tsãtǻ 1A	mĭdzyï 1B
Latin etymon	*CANTARE	*MANDUCARE
Gloss	'sing'	'eat'
1SG	tsãt-ŏ	mïdz-ŏ
2SG	tsãt-e	mïdz-e
3SG	tsãt-e	mïdz-e
1PL	tsãt-ẽ	mïdz-ẽ
2PL	tsãt-ǻ	mïdz-yï
3PL	tsãt-õ	mïdz-õ
IMPERFECT		
1SG	tsãt-ǻò	mĭdz-īyŏ
2SG	tsãt-ǻe	mĭdz-īye
3SG	tsãt-ǻe	mĭdz-īye
1PL	tsãt-ǻẽ	mĭdz-īyẽ
2PL	tsãt-ǻī	mĭdz-īyi
3PL	tsãt-ǻõ	mĭdz-īyõ

Table 54. Class 1a and 1b in Francoprovençal in Bagnes (Bjerrome 1957:86-88)

One should note that, although the vocalic development in palatal environments yielded a new thematic vowel /i/ this development does not amount to heteroclisis with reflexes of Latin fourth conjugation presenting a thematic vowel /i/. Some inflections are identical (in particular in the imperfect indicative), but other developments are peculiar to the new class, notably the differential treatment of second person plural in the present indicative. The result is a system of inflectional classes that functions in terms of defaults, with class 1A and 1B being subtypes of each other in an inheritance hierarchy model. In Bagnes for example (Bjerrome 1957), the future and subjunctive have similar realizations for all classes, class 1B patterns with class 1A for present indicative (except second person plural) and conditional, and patterns with class 2 for imperfect indicative. The second person indicative present, infinitive and past participle are peculiar to class 1B.

Francoprovençal thus shows the creation of a new inflectional class distinction following two sound changes which involve the affix boundary. First, the final consonant of the root palatalizes; then it triggers a vowel change in the thematic vowel.

4.1.2.2. Romanian

In Romanian (Romance, Indo-European; Maiden 2016), and particularly in Istro-Romanian varieties (Maiden 2017), reflexes of verbs belonging originally to the Latin fourth conjugation

underwent conditioned sound change, with the result that an additional fifth conjugation was created (Maiden 2017:250-251; Maiden 2016:100), although typically for a Romance language, Romanian only inherited four conjugations.

In standard Romanian, the class of verbs continuing the Latin fourth conjugation whose root ended in *-[rr] underwent centralization of the thematic vowel, whereby [i] centralized to [ɨ] (orthographically both <â> and <î>, as in *omorî* and *omorând* in Table 55) and [e] centralized to [ə] (orthographically <ă>). Later the phonological conditioning of that vowel centralization was lost with the merger of [rr] with [r], leading to a situation in which some verbs with stem-final /r/ had the centralized vowel and others did not: the distribution became opaque. A contrasting example of a verb without centralization is *sări* 'jump', as opposed to a verb with a centralized thematic vowel such as *omorî*. Note that these cases sometimes created heteroclisis, some of the new endings with the new, centralized thematic vowel corresponding to identical endings in the first class of verbs (Maiden 2017:251).

	Class 4a *fugi* 'run away'	Class 5a *omorî* 'kill'
Etymon	*FUGI:RE	Slavic *UMORITI
INFINITIVE	fugi	omorî
GERUND	fugind	omorând
PAST PARTICIPLE	fugit	omorât
PRESENT		
1SG	fug	omor
2SG	fugi	omori
3SG	fuge	omoară
1PL	fugim	omorâm
2PL	fugiţi	omorâţi
3PL	fug	omoară

Table 55. Fourth and fifth conjugations in Romanian (Maiden 2017:251)

4.1.2.3. Istro-Romanian

Istro-Romanian shows a different example of a new class originally arising from fourth conjugation verbs (Maiden 2017). In Istro-Romanian, a new class is developed with a new thematic vowel [ɛi], which has exactly the same paradigmatic distribution as the thematic vowel [i] of the fourth conjugation (Table 56 and 57). Both thematic vowels are present in first and second person

plural in the present, and in the imperfect and conditional, as well as in the non finite forms infinitive, past participle and gerund. Both classes also present the augment characteristic of the fourth conjugation in the Romance languages, in the present singular and 3PL (*-es*, *-ef*, etc).

	1SG	2SG	3SG	1PL	2PL	3PL
Present	ko'ses	ko'seʃ	ko'sɛ	ko'**s**im	ko'**s**its	ko'ses
Imperfect	ko'**s**i̯am	ko'**s**i̯ai̯	ko'**s**i̯a	ko'**s**i̯an	ko'**s**i̯ats	ko'**s**i̯a
Conditional	ko'**s**ir	ko'**s**iri	ko'**s**ire	ko'**s**irem	ko'**s**irets	ko'**s**iru

Table 56. Fourth conjugation verb ko'si *'to mow' in Istro-Romanian (Maiden 2017:252)*

	1SG	2SG	3SG	1PL	2PL	3PL
Present	ko'pes	ko'peʃ	ko'pɛ	ko'p**ɛ**i̯m	ko'p**ɛ**i̯ts	ko'pes
Imperfect	ko'p**ɛ**i̯am	ko'p**ɛ**i̯ai̯	ko'p**ɛ**i̯a	ko'p**ɛ**i̯an	ko'p**ɛ**i̯ats	ko'p**ɛ**i̯a
Conditional	ko'p**ɛ**i̯r	ko'p**ɛ**i̯ri	ko'p**ɛ**i̯re	ko'p**ɛ**i̯rem	ko'p**ɛ**i̯rets	ko'p**ɛ**i̯ru

Table 57. Fifth conjugation verb ko'pɛi̯ *'to dig' in Istro-Romanian (Maiden 2017:252)*

The identical distribution of the thematic vowel and the presence and identical distribution of the augment argue in favour of a common origin for the two classes, that is in fact for the fifth class to have originated in the fourth conjugation, with some sound change. The fifth conjugation is also peculiar in that all its members are borrowings from Italian or Croatian. It is thus a kind of etymological class (see Chapter 3). Fourth class is normally the productive class where borrowings are adapted in Istro-Romanian. The new thematic element [ɛi] would have originated from an initial reanalysis of root and affix boundary. All verbs in that class have been borrowed from Italian or Croatian with a root final *-a*, and included in the productive fourth conjugation with a thematic vowel [i]. Later on, the presence of the thematic vowel is suggested to have triggered raising of root final *-a* before thematic [i] to [ɛ], and the resegmentation of the two vowels as a diphthong. This diphthong being present in all cells where the fourth conjugation thematic vowel is present, it was reinterpreted as the new thematic element for that class (Maiden 2017).

4.1.2.4. Old French

In the passage from Latin to Old French, tonic free mid-vowels diphthongized, in two waves. The first wave affected tonic free mid-low vowels and is common to a number of Romance languages, including Spanish and Italian. Basically, mid-low vowels [ɛ] and [ɔ] diphthongized to [jɛ]

and [wɔ] (Zink 1986:52-56)[33]. A second wave of diphthongization was largely confined to Old French: tonic free mid-high vowels [e] and [o] diphthongized to [ej] and [ow] (Zink 1986:57-60).

The process of diphthongization affected in particular a number of verb roots. The specific conditioning (vowels had to be stressed and in an open syllable to diphthongize) meant that this created a number of specific alternations in verbal stems, following patterns of stress and syllable structure. This affected in particular a number of Latin first conjugation verbs, which meant the creation of a number of inflectional subclasses dividing the first conjugation between those verbs which did not present diphthongizations, and four classes of verbs presenting various new stem alternations, as shown in Table 58 (for the present indicative only).

INF	chanter	esperer	lever	plorer	trover
GLOSS	sing	hope	hold	cry	find
1SG	chant	espeire	lief	plour	truef
2SG	chantes	espeires	lieves	ploures	trueves
3SG	chante	espeire	lieve	ploure	trueve
1PL	chantons	esperons	levons	plorons	trovons
2PL	chantez	esperez	levez	plorez	trovez
3PL	chantent	espeirent	lievent	plourent	truevent

Table 58. Old French verb classes from the Latin first conjugation following diphthongization (after Zink 1986; Fouché 1931; Picoche 1979)

Note that diphthongization also affected verbs in other classes, which meant the creation of a number of other new subclasses, all following an N-pattern distribution of stem alternants, which is a Romance specific pattern of distribution of root allomorphy opposing the three persons in the singular and the third person plural present tense with one alternant to the rest of the paradigm with another alternant (Maiden 2009 among other works). This stem distribution was also found in other types of verbs with alternations not due to diphthongization (ex. *parler* 'speak'). Some of the inflectional classes based on N-pattern stem alternations due to diphthongization were later regularized to the first class model *chanter* 'sing', by analogy (Esher 2017).

33 There is some debate as to whether the diphthongs were originally rising or falling; in any case the result in Old French is a rising diphthong.

4.1.3 Metaphony

Metaphony is the term used in Romance linguistics as an equivalent of German *Umlaut* (Calabrese 2011). It consists in the raising (and in some cases diphthongization) of a stressed vowel which is triggered by the presence of a following final high vowel. As Schmid (2016) puts it, "metaphony is triggered by a high vowel /i/ or /u/ occurring in a following syllable (usually but not always adjacent), which determines the raising or diphthongization of preceding vowels". As such, it is a form of regressive vowel harmony. But contrary to usual types of vowel harmony, it has as a specific target a stressed vowel. In many southern Italo-Romance varieties, this has given rise to a number of morphological alternations which are unmotivated in synchrony, following the reduction of final unstressed vowels.

An example of metaphonic alternations still conditioned in synchrony by the presence of a [+high] segment in a post-tonic position is found in the dialect of Servigliano (Maiden 1991:160-161; Calabrese 2011). In this dialect, stressed mid vowels are raised in the presence of a post-tonic high vowel /i/ or /u/. Thus /e/ and /o/ are raised to [i] and [u], and stressed /ɛ/ and /ɔ/ are raised to [e] and [o]:

Singular	Plural	Meaning
'fjore	'fjuri	flower
'botte	'butti	barrel
'pɛde	'pedi	foot
'dɛnte	'denti	tooth

Table 59. Metaphony in Servigliano (Calabrese 2011)

But in a number of dialects, metaphony can no longer be justified by synchronic phonological rules, following either the loss or the reduction of final unstressed vowels. In such varieties, metaphonic alternations have fully morphologized, and now realize morphological alternations. In the dialect of Arpino for example, final unstressed vowels normally reduce to schwa (except for final /a/). This means that previous metaphonic alternations are now used as a morphological realization of morphosyntactic features in nominal and adjectival paradigms. Thus in adjectival paradigms of the class I/II, feminine plural and masculine plural are only distinguished by the stem

vocalic alternation (Table 60).

MASC.SG	FEM.SG	MASC.PL	FEM.PL	Meaning
'sulə	'sola	'sulə	'solə	alone
'nirə	'nera	'nirə	'nerə	black
'bwonə	'bɔna	'bwonə	'bɔnə	good
'vjekkjə	'vɛkkja	'vjekkjə	'vɛkkjə	old

Table 60. Class I/II adjectival metaphonic alternations in Arpino (Maiden 1991:165; Calabrese 2011)

The interesting fact in terms of inflectional class creation is that all those words originally belonged to the same inflectional class. After a number of sound changes have occurred (first metaphony, then final vowel neutralization), they now belong to four different subclasses: all of those are characterized by a marker -a for the feminine singular, and by syncretism between masculine singular and plural. But the actual realization of the distinctions being in part or in totality expressed by vocalic stem alternations, they now belong to different classes as those alternations differ: [o/u] for the word meaning 'alone', [e/i] for the word meaning 'black', and so on.

The same thing holds for nouns originally in class III (Table 61). Thus nouns originally belonging to the same inflectional class are now members of three different inflectional classes following two sound changes, first metaphony, then the neutralization of final unstressed vowels.

Singular	Plural	Meaning
'fjorə	'fjurə	flower
'mesə	'misə	month
'vɛrmə	vjermə	worm

Table 61. Class III nouns in Arpino (Maiden 1991:165; Calabrese 2011)

Similar kinds of morphologized alternations for nominal paradigms can be found in a number of Italian dialects (see Maiden 1991:162-170), and in German nouns.

4.2. A theoretical problem: when do inflectional classes appear under phonologically conditioned allomorphy?

The sound changes examined hitherto, in particular cases of metaphony, raise the theoretical

question of when one can consider that a given sound alternation has morphologized. The traditional view is to consider that they are part of the morphology only when they are not motivated phonologically in synchrony, that is when one can be sure that only a morphological motivation persists for such alternations. This is for example the case for metaphony in the dialect of Arpino, as shown in section 4.1.3., where in synchrony the metaphonic alternations can no longer be said to be triggered by the presence of a final high vowel. In such cases, there is proof that the vocalic stem alternation has already morphologized at some point earlier, but no indication as to when exactly it morphologized.

But a range of evidence shows that the morphologization of low level phonetic phenomena occurs before the loss of the conditioning environment. There is first a logical argument. If metaphony were still a synchronic phenomenon, triggered by the presence of a final high vowel, then the loss of the conditioning environment (the final high vowel) should mean the loss of the metaphonic alternation. What one should find in Arpino would be the following hypothetical scenario: metaphonic raising due to the presence of a final high vowel > neutralization of the final vowel, which removes the conditioning environment for the metaphonic raising, and thus > absence of the metaphonic raising. The metaphonic alternation must thus have morphologized earlier than the neutralization of final vowels for the situation actually found in Arpino to occur, with a scenario on these lines: metaphonic raising due to the presence of a final high vowel > morphologization of the distinction > neutralization of final high vowels. It is thus highly probable that even when the conditioning environment is still present, the metaphonic vowel alternation is already morphologized. An alternation may appear to be motivated phonetically in synchrony but it can also be recruited by the morphology to signal (or co-signal) some feature value.

This is essentially the line of argument adduced in Maiden (1991) for Italian metaphony. Maiden states that in cases of metaphony, "phonetic and morphological conditioning *coexist*; also, metaphonic alternation is probably directly associated with the morphosyntactic category, independently of the inflectional suffix" (Maiden 1991:3). To identify instances of

morphologization of metaphony, Maiden uses in particular evidence from analogical extension of the results of the sound change to environments that would not normally trigger it (Maiden 1991:12). One of the arguments for the morphologization of metaphony comes from the differential treatment it receives in diachrony in verbal and nominal paradigms under neutralization of final vowels in some varieties. Thus in the variety of Cesena, metaphony is neutralized in verbal environments, but preserved in all nominal paradigms, which shows that it has morphologized as an exponent of number in nominal paradigms (Maiden 1991:194-195). Similarly, many Romagnol varieties present, in addition to masculine plural suffix *-i*, a new feminine plural suffix *-i*, which does not trigger metaphony (Maiden 1991:195). This is proof that the metaphonic alternation has morphologized, and that it is not only a function of the presence of a given final high segment in the word.

The fact that metaphony is at the same time a phonetic and a morphological phenomenon can be proved, according to Maiden (1991:218-222) by a number of facts: the creation of metaphonic gender alternations analogically outside the presence of a high vowel; the retention of metaphony in Serviglianese when neuter nouns adopt inflection *-o* instead of *-u*; the developments of Neapolitan class one feminines; and the "mixed conditioning configurations" whereby "metaphonic assimilation continues to operate before final /i/ despite neutralization of the masculine singular inflection -u and consequent morphologization of gender alternation in class one singulars" (Maiden 1991:218). The main argument is that the phonetic process is morphologized so early that it is sensitive to the morphosyntactic environment in which it is operating: specifically the phonetic process affects a wider range of input vowels, and produces a greater degree of phonetic assimilation, in the verb rather than anywhere else. This is evidence for the inseparability of phonological processes and morphologization.

A number of studies confirm this recruitment of low level phonetic variation by the inflectional morphology of languages, showing that subphonemic distinctions can, and in fact generally do morphologize. Baayen et al. (2003) show that frequency plays a role in the recognition of plural

word forms in Dutch. They conclude that even regular plural word forms are likely to be stored. In cases where the plural appears to be simply the concatenation of a singular stem and a plural suffix, they show that at the phonetic level this is not the case. The singular stem is in particular shown to have longer duration than the phonologically identical plural stem. They state that "at a given speaking rate, therefore, speakers tend to take more time to produce singular forms than they do to say a stem embedded within a plural. These differences should encourage the development of plural access representations. If the plural in the spoken input mismatches with the stored phonological form of the singular, recognition via morphological parsing will be slower, and plurals will be more likely to develop independent representations and thus to be recognized via full-form representations" (Baayen et al. 2003:364). Their study is an argument in favour of the use of low level phonetic differences in the word recognition of inflected word forms. The plural is not just a concatenation of the singular plus a plural affix, but it has fine grained differences in the representation, storage and production of the stem.

Similar effects are found in follow up studies (Kemps et al. 2005a; Kemps et al. 2005b; see also Blevins 2016b:54-55). These confirm in particular that a stem found in isolation as a singular form has a longer duration than the arguably phonologically identical stem found as the stem for forming a plural by suffixation. It confirms that speakers make use of fine grained phonetic differences, in particular at prosodic structure, in order to distinguish such stems.

What all these studies mean for the question of morphologization of phonetic phenomena is that phonetic cues are normally always part of the morphological retrieval of forms, and thus that a phonetic alternation is very likely to be recruited by the morphology as a signans of morphosyntactic categories. One should thus consider that in the case of metaphony, inflectional classes are created very early in the presence of such alternations.

4.3. Sound change and the rise of inflectional classes in South Halmahera-West New Guinea languages

This section proposes an analysis of the development of verb inflection in South Halmahera-West New Guinea languages, particularly those belonging to the Biak-Yapen group, which show the creation of inflectional classes due to various sound changes, including the sound adaptation of affix boundaries under grammaticalization, and metaphony.

4.3.1 Grammaticalization of bound pronouns and boundary sound adaptation

Verbal paradigms in South Halmahera-West New Guinea languages inflect for subject agreement via the prefixation of subject bound pronouns. Such pronominal affixes can be considered as either agreement markers or as incorporated pronouns at the level of the syntax, but in any case they are clearly inflectional at the level of morphology (see Corbett 2006:99-112). Some languages also inflect for object agreement. This section will only consider subject agreement marking. Subject agreement markers are clearly linked diachronically to the free corresponding personal pronouns, as can be seen for Wooi in Table 62 and for Biak in Table 63:

	Free pronoun	Bound pronoun
1SG	ya	y-
2SG	aw	bu-
3SG	i	ti-
1DU.EXCL	aru	u-
1DU.INCL	taru	tu-
2DU	maru	mu-
3DU	haru	hu-
1PL.EXCL	ama	ma-
1PL.INCL	tata	ta-
2PL	mia	me-
3PL	hia	he-

Table 62. Free and bound pronouns in Wooi (Sawaki 2016:75-78)

	Free pronoun	Bound pronoun[34]
1SG	aya	ya-
2SG	au	wa-
3SG	i	i-
1DU.EXCL	nu	nu-
1DU.INCL	ku	ku-
2DU	mu	mu-
3DU	su	su-
3PCL	sko	sko-
1PL.EXCL	(i)nko	(i)nko-
1PL.INCL	ko	ko-
2PL	mko	mko-
3PL.ANIM	si	si-
3PL.INAN	na	na-

Table 63. Free and bound pronoun in Biak (Mofu 2005:30)

As all South Halmahera-West New Guinea languages show such bound pronominals, one can reconstruct that they grammaticalized in proto-South Halmahera-West New Guinea, at least in the form of clitics. Some South Halmahera-West New Guinea languages also present object enclitic pronouns. Subject pronouns were first used obligatorily adjacent to the verb, in a preverbal position congruent with the SVO preferred word order of these languages, before cliticizing and eventually affixing to the verb root. That pronouns grammaticalized can be shown by the fact that they are obligatory (Examples 1 and 2), and that they co-occur with an expressed subject noun phrase (Example 3). The following examples illustrate these facts for Biak:

1) ya-mbran be mnu
   1SG-walk to village
   'I walked to the village' (Mofu 2005:71)

2) *mbran be mnu
   walk to village
   'I walked to the village' (Mofu 2005:71)

3) mansar ine i-duf kaku
   old.man this 3SG-be.sick very

34 The bound pronoun form here is given for class 1 verbs.

'This old man is very sick' (Mofu 2005:83)

As usual in such processes of grammaticalization, a certain level of phonetic erosion occurred, depending in particular on the phonological properties of the root they attached to. This was in particular the case where the verb root started in a consonant or in a vowel. Such sound adaptation gave rise to two major inflectional classes in most of the languages of the family, conditioned by the phonological shape of the root.

Such examples show the creation of inflectional classes not only through sound change, but through the combined factors of grammaticalization and sound adaptation at morph boundary.

4.3.2 Metathesis

The process of metathesis and its implications for morphological paradigms, notably the development of verbal infixation, has been studied at length by Gasser (2015). It is attested in a number of South Halmahera-West New Guinea languages belonging roughly to a proto-Biak-Yapen group comprising all the South Halmahera-West New Guinea languages spoken on Biak and Yapen islands in addition to Yaur, Yeretuar, and Iresim (Gasser 2015:4). Metathesis must thus have occurred in the protoform of that language group (Gasser 2015:6). Table 64 shows the singular realizations of subject agreement on verbs in the languages which show infixation in the second and third person singular.

	1SG		2SG		3SG	
Language	PRO	AGR	PRO	AGR	PRO	AGR
Wooi	yau	y- / i-	au	bu- / -u-	i	ty- / -i-
Munggui	yau	y- / i-	au	w- / -u-	i	ty- / -i-
Pom	yau	y- / i-	au	w- / -u-	i	di- / -i-
Papuma	yau	y- / e-	au	w- / -u-	i	t- / -i-
Busami	yau	ya-	au	w- / -u-	i	s- / -i-
Wamesa	yau	y- / i-	au	bu- / -u-	i	di- / -i-
Ansus	yau	y- / e-	au	bu- / -u-	i	d- / -i-
Serui	yau	y- / i-	wau	bu- / -u-	i	d- / -i-
Ambai	yau	y- / i-	wau	bu- / -u-	i	d- / -i-
Wabo	aya	ay- / a-	awa	b- / -o-	i	d- / -i-
Kurudu	aya	ay- / a-	awa	b- / -u-	i	d- / -i-
Biak	aya	y- / ya-	au	w- / -u-	i	d- / -i-
Waropen	ya	y- / ya	auo	au- / a-	i	i(y)- / -i-
Roon	ya	y(a)- / i	aw	w(a)- / -u-	(t)i	t- / -i-

Table 64. Infixing realizations in Biak-Yapen languages (after Gasser 2015:5)

Gasser (2015) describes the process leading to infixation in some verbs of the Biak-Yapen group as coming from Perceptual Metathesis (Blevins & Garrett 1998). Gasser notes that perceptual metathesis "occurs when the acoustic cues associated with a given segment have a particularly long duration. This persistence creates ambiguity regarding the origin of the cues in the word, leading to reinterpretation of the origin of the cues in question in a non-historical position" (Gasser 2015:7). This type of metathesis occurs in particular for high vowels, as shown by attestations in Greek and in some Bantu languages (Blevins & Garrett 1998). In Proto-Biak-Yapen, both labiality and palatalization were associated with perceptual metathesis for two different affixes, the 2SG and 3SG affixes, in a morph specific process. A number of studies show that the timing of segments is different when morph internal and when at morph boundaries. In particular, timing seems to be specified morph internally, but there is huge variation at morph boundary (Cho 1998, 2001). This non-specification of timing may allow for coarticulation and reduced timing at morph boundary. Gasser synthesizes the development of infixes as follows: "The first step towards metathesis was increased coarticulation of the high vowel with the following consonant, leading to rounding or palatalization of that consonant. As coarticulation continued to increase, the secondary articulations of the consonant developed into a full offglide. In these forms, the segmental source of the palatalization or labialization was ambiguous, and was reanalysed as originating solely after the consonant." (Gasser 2015:9), that is 'metanalysis' in Ohala's terms. Thus 2SG *au* saw the labio-velar element infixing,[35] first with apical-initial roots, and SG *i* saw the palatal element infixing first in velar-initial roots. There was later some extension of those infixing elements to the same set of roots. In Wamesa, this extension went further as to become the sole pattern for consonant initial roots. In Biak, only some of the consonant initial roots follow this pattern. A number of verbs, such as **ra* 'go' follow this pattern in most Biak-Yapen languages.

35 Gasser posits an affix *bu-* in proto-Biak-Yapen, because it is the most attested reflex in those languages. But SHWNG had *au*, and some of the infixing languages have *wa-* as in Biakic languages. I posit that the initial element was *au* and that further change occurred in proto-Yapen yelding *bu-* as the second singular prefix.

Metathesis is a phonetic phenomenon which gets morphologized in Biak-Yapen languages, much in a similar fashion as metaphony did in Italian dialects. Again, low level phonetic phenomena matter for the morphology, and can get morphologized.

4.3.3 Further changes in Biak

A number of further changes affected the infixal class of verbs in Biak. Some of the changes have to do with some sound adaptation, other changes are due to analogy. Only the changes due to sound change are treated in this section.

There is a further metathesis which has not been accounted for until now in those paradigms, that of the second person singular marker.

The further changes in Biak all concern the marking of the second singular in infixing verbs. Depending on the initial consonant of the stem, the infixal marker *-w-* does not surface or surfaces as a change in the stem consonant. The marker does not surface in verbs with an initial nasal bilabial consonant /m/ or with an initial /f/, as shown in Table 65 for the verbs *mam* 'see' and *fas* 'write' (only the singular paradigm is shown). This forms a subclass of the infixal class which is conditioned phonologically. Verbs starting with a bilabial voiced approximant /β/ (orthographically <v>) show a change in the initial consonant of the stem strengthened to /b/, as shown in Table 65 for the verb *vov* 'sell'.

	so	mam	fas	vov
	throw	see	write	sell
1SG	ya-so	ya-mam	ya-fas	ya-vov
2SG	s-<w>-o	mam	fas	bov
3SG	s-<y>-o	m-<y>-am	f-<y>-as	v-<y>-ov

Table 65. Infixal subclasses in Biak (Steinhauer 2003:8-9)

4.3.4 Further changes in Wooi

Wooi shows further creation of inflectional classes following a number of sound changes and analogy. The data in this section are taken from Sawaki (2016), but with a different analysis. Sawaki accounts for allomorphy in terms of underlying structure being identical for all classes, with

a number of phonological adaptations that are specific to each affix for accounting for allomorphy. I do not posit such identical underlying structure, but work directly with surface structures which form a number of different inflectional classes.

Wooi first extended the infixal pattern to all consonant-initial verbs, as seems to be the case for all Yapen languages: the data are scarce[36], but the fact that Wamesa also shows an infixal pattern for all consonant-initial verbs argues in favour of a proto-Yapen, or at least a proto-West Yapen phenomenon. In Wooi, two further sound changes affected some of the affixal realizations to yield new inflectional classes, mainly palatalization and vowel cluster simplification. These changes were local and did not affect all roots: they only affected those roots with the specific phonological environment (either consonant or vowel-initial). The present system of inflectional classes in Wooi is as shown in Table 66.

	V-initial		C-initial		
	-ena	-avayang	-ra	-ha	-kavio
	'sleep'	'buy'	'go'	'call'	'talk'
1SG	y-ena	y-avayang	ra	ha	kavio
2SG	bu-ena	b-ovayang	r<u>a	h<u>a	k<o>vio
3SG	c-ena	t-evayang	r<i>a	h<i>a	k<e>vio
1DU.EXCL	ur-ena	ur-avayang	un-da	u-sa	ung-kavio
1DU.INCL	tur-ena	tur-avayang	tun-da	tu-sa	tung-kavio
2DU	mur-ena	mur-avayang	mun-da	mu-sa	mung-kavio
3DU	hur-ena	hur-avayang	hun-da	hu-sa	hung-kavio
1PL.EXCL	mat-ena	mat-avayang	man-da	ma-sa	mang-kavio
1PL.INCL	tat-ena	tat-avayang	tan-da	ta-sa	tang-kavio
2PL	met-ena	met-avayang	men-da	me-sa	meng-kavio
3PL	het-ena	het-avayang	hen-da	he-sa	heng-kavio

Table 66. Inflectional classes in Wooi (Sawaki 2016:170-178)

It seems that initially verbs whose first vowel was /a/ underwent a simplification of the hiatus, as shown in Table 66 for *avayang* 'buy' and *kavio* 'talk'. In each case the expected sequence *-ua-* for second person singular was reduced to *-o-*, and the expected sequence *-ia-* for third person singular was reduced to *-e-*. The result is that for consonant-initial verbs whose first vowel is an /a/, singular persons are now distinguished only by vowel alternations on the stem. The same happens for some

36 In the limited fieldwork I did on Pom, Ansus, and Munggui (a dozen verbs in total), all West Yapen languages, all consonant-initial verbs were infixing.

vowel initial verbs. Now it is not possible to posit that this was entirely due to phonological factors: a number of verbs whose first vowel is /a/ did not undergo this change. This is in particular the case for *ang* 'eat', which inflects on the same pattern as *ena* 'sleep', and *mata* 'itch' and *tawa* 'fall' which inflect like *ra* 'go' (provided the allomorphy of the plural forms is accounted for phonologically for *ra*).

Palatalization and vowel cluster simplifications seem to be ordered in diachrony. For the results we see in particular in 3SG affixes, vowel cluster simplification must have occurred first. This leads to positing a simple palatalization of the affix *ti-* for third person singular in vowel-initial verbs: *ti-* palatalized to *c-* (phonetically [tʃ]). The vowel simplification and raising yelded two new subclasses, while the palatalization only reinforced the identity of one class with regard to the other vowel-initial class. Such reinforcement is a common output of sound change, which does not always create new classes, as shown in the next section.

4.3.5 Further changes in Biak Betew

Biak Betew is the dialect of Biak spoken in the western part of the Raja Ampat, the westernmost large settlement of Biak speakers traditionally, other than Ternate[37]. The dialect is quite distinctive in that it has undergone some sound changes, as well as some distinct lexical innovations. One sound change has important bearings on the inflectional class system, and is as follows:

/s/ --> [h] in front of /n/

/s/ --> [s] in front of /i/ and /j/ (high front vowel and glide)

/s/ --> [t] elsewhere[38]

As 3PL.AN is marked either with *si-* for class 1, or with *s-* for other classes, this sound change affects the form of the inflectional marker for this cell, as shown in Table 67:

37 All the data on Biak Betew come from fieldwork with one speaker of the dialect, David Dimara, from Mutus island.

38 This is a typologically unusual change (the reverse change is more commonly attested), but it can be ascertained from comparative evidence. Proto South Halmahera-West New Guinea *t merged with *k in all Biakic dialects, which nowadays do not present a phoneme *t. Only in Western varieties such as Biak Betew did this sound reappear, and systematically corresponds to /s/ in other Biak dialects.

	PRO	Class 1	Class 2	Class 3
		'wake up'	'come'	'sleep'
1SG	aya	yakwoek	yarama	yenef
2SG	au	wakwoek	rwama	wenef
3SG	i	ikwoek	ryama	denef
1DU.INCL	ku	kukwoek	kurama	kuyenef
1DU.EXCL	nu	nukwoek	nurama	nuyenef
2DU	mu	mukwoek	murama	muyenef
3DU	tu	tukwoek	turama	tuyenef
3PAUC	tko	tkokwoek	tkorama	tkenef
1PL.INCL	ko	kokwoek	korama	kenef
1PL.EXCL	nko	nkokwoek	nkorama	nkenef
2PL	mko	mkokwoek	mkorama	mkenef
3PL.ANIM	si	sikwoek	trama	tenef
3PL.INAN	na	nakwoek	ndrama	nenef

Table 67. Verb classes in Biak Betew

It could still be considered a simple case of conditioned allomorphy if two elements did not present evidence for both a phonologization of the distinction s/t, and a morphologization of the distinction in terms of classes. In terms of phonologization of the distinction s/t, there is a widespread phenomenon in the variety I examined of glide reduction for postconsonantal high front glides[39]. Thus, the phonological environment conditioning the allophone [s] being lost, it can be considered to have achieved phonemic status, as shown in the following examples:

	Biak	Biak Betew
'dry'	syor	sor
'boil'	son	ton
'bow'	syom	som

Table 68. Glide reduction in Biak Betew

The second interesting point is that one of the markers is now extended to a different environment. Below are partial paradigms for three verbs belonging to class 3, as they are vowel initial roots:

39 This might not affect all postconsonantal environments in the same way (it doesn't seem to get reduced after /r/ for example), but certainly affects /s/.

	'get up into your arms'	'sleep'	'drink'
ROOT	awen	enef	inem
1SG	yawen	yenef	yinem
2SG	wawen	wenef	winem
3SG	dawen	denef	dinem
1PL.INCL	kawen	kenef	kinem
1PL.EXCL	nkawen	nkenef	nkinem
2PL	mkawen	mkenef	mkinem
3PL.ANIM	tawen	tenef	tinem
3PL.INAN	/	nenef	ninem

Table 69. Three vowel-initial verbs in Biak Betew

If the phonological conditioning on allomorphs of /s/ were still operating, one would expect a form **sinem* for 3PL.AN of *inem* 'drink'. Instead, what one sees is an extension of the marker *t-* to all members of class 2 and 3 as a marker of 3PL.AN, while *si-* is restricted to marking class 1. It is, in fact, the creation of a default form by extending the pattern analogically.

While these changes do not directly affect the inflectional class system of Biak Betew, which has still the same number of classes, these classes are strengthened by markers that are more specific to a given class, and further away phonologically. In particular, the markers of 3PL.ANIM are no longer conditioned phonologically.

4.4. Conclusion

Sound change has long been recognized as a main source of inflectional classes. Some of the sound changes involving the creation of new classes arise in existing paradigms, and create subclasses in an inheritance hierarchy model. Other sound changes apply when grammaticalization processes occur, but adapting the sounds that are present at affix boundaries. I now turn to this second main source of inflectional classes, grammaticalization.

5. Grammaticalization

This chapter is devoted to instances of creation of inflectional classes due to grammaticalization. Some types of grammaticalization have been recognized as sources for inflectional classes (see Dammel 2011), mostly the grammaticalization of new affixal distinctions with some sound adaptation at morph boundaries, as treated in the preceding chapter. This chapter treats two new scenarios which have not yet been recognized as potential sources of inflectional classes, the grammaticalization of auxiliaries, and cycles of grammaticalization of bound pronominals.

Grammaticalization is a type of linguistic change whereby lexical elements become more grammatical in nature, and free forms become bound. It thus presents two different clines: a functional cline from lexical to grammatical with in-between stages; and a formal cline from free forms to affixes through various stages of cliticization. The two clines need not be in an exact parallel: a lexical element can become more grammatical without becoming an affix, but it will show reduced distributional properties, and potentially a reduced phonological form. Grammaticalization occurs in discrete steps, that can be reconstructed for a number of constructions along the formal and the functional clines.

The fact that inflectional classes can be created from existing affixes (i.e. from affixes that have already grammaticalized from another source) is well recognized (Dammel 2011; Harvey 2008). Affixation and the rise of inflectional classes through affixation can be said to be an endpoint of grammaticalization processes. After that, it is often said that the only possibility for an affix to evolve is through loss (see Norde 2002). If a marker with meaning cliticizes and then affixes to a root, it can undergo further semantic bleaching and lose its meaning completely. But it can also retain some morphological meaning if it is associated with a subpart of the lexicon. Thus, in Guugu Yimidhir there are inflectional class markers which can be reconstructed to have been originally TAM markers on the verb (Harvey 2008; others have analysed them as coming from reanalysis, see Chapter 6). TAM markers are in themselves highly grammaticalized elements, which have gone all

the way through the affixation cline from free elements to affixes through a stage of cliticization. Harvey shows that these elements were semantically bleached further, leaving only a formal element associated with a subset of verbs, which created inflectional classes in the language.

Another possibility is the creation of inflectional classes because of phonological adaptation of free elements to the phonology of the root when a process of cliticization occurs. This is what can happen when pronouns grammaticalize as agreement markers on verbs (see Chapter 4, and section 5.2. below). Sometimes, a differential grammaticalization of pronominal elements with different roots will create inflectional classes of a very specific type: inflectional class by position class. I show that this is what happened in Arapesh (section 5.2.3), which is a novel source for inflectional classes. Finally the grammaticalization of auxiliaries has not been until now recognized as a source for new inflectional classes (section 5.1.).

5.1. Auxiliation

One way to acquire inflectional class distinctions through grammaticalization is the creation of an auxiliary class. Either different lexemes grammaticalize different auxiliaries (Romance, and to some extent Basque), or some verbs inflect in a synthetic way while others inflect periphrastically through the use of an auxiliary (Basque). This evolution can in principle be available to all cases where a periphrastic construction starts to compete with a synthetic construction, which can give rise to two competing classes of exponence (such as the possessive paradigm of Maltese nouns), but I only treat the example of auxiliaries here. More eamples coming from possessive inflection are discussed in Chapter 7.

5.1.1 Romance auxiliary selection as segregated inflectional class

Romance languages show periphrasis in the part of their verbal paradigm associated at least originally with perfective aspect. Compound tenses include present/past perfect and pluperfect indicative, as well as past and pluperfect subjunctive, and future perfect. Inflectional periphrasis can

be considered as inflectional morphology when the periphrastic expression occupies a cell in the paradigm of a lexeme and enters into paradigmatic oppositions with synthetic forms (Spencer 2003; Bonami 2015), as is the case for Romance perfective tenses.

Inflectional periphrasis has attracted much attention in recent years (Spencer 2003; Ackerman &Stump 2004; Bonami & Samvelian 2009; Brown & al. 2012; Bonami 2015 among others). Brown & al. (2012) in particular set out to define inflectional periphrasis in the framework of Canonical Typology. They consider inflectional periphrasis to be at the same time morphology and syntax. Specific constructions can be mapped on a cline from more morphological to more syntactic. They first define inflectional periphrasis informally as "the situation where we find two (or more) words even though we had a reasonable, morphology-based, expectation of finding only one" (Brown & al 2012:237). The morphology-based expectation of a single form comes from the paradigmatic distribution of forms. In a system where tense, aspect and mood are opposed on verbs for example, one can expect synthetic forms for each of the oppositions. If one finds that an opposition is rendered by a multi-word expression, this is an instance of inflectional periphrasis. More formally, the periphrasis is the expression of feature intersection that defines a cell in a paradigm. Thus in the case of Romance perfective tenses, only perfective tenses show a periphrasis, where the system of imperfective tenses would lead one to expect a single form per cell. Inflectional periphrasis in that sense can be said to be morphology, in that a multi-word expression occupies a cell in an inflectional paradigm. In addition to occupying a cell in the paradigm, Ackerman & Stump (2004) argue that non-compositionality of the periphrasis is also a key defining element, sufficient to define periphrasis. Brown & al. (2012) on the contrary argue that it is a necessary component, but not sufficient, as idioms are also non-compositional. Thus, the perfective periphrasis in Romance expresses an aspectual distinction of perfectivity, but none of its component elements (the past participle, the auxiliary) is inherently perfective. Perfectivity is thus non-compositional in such expressions.

Table 70 shows the present and the perfect sub-paradigm of the verbs *aller* 'go' and *manger* 'eat'

in French, which each select a different auxiliary, *être* 'be' for *aller*, and *avoir* 'have' for *manger*:

		aller 'go'	manger 'eat'
PRESENT	1SG	je vais	je mange
	2SG	tu vas	tu manges
	3SG	il va	il/elle mange
	1PL	nous allons	nous mangeons
	2PL	vous allez	vous mangez
	3PL	ils vont	ils/elles mangent
PERFECT	1SG	je suis allé(e)	j'ai mangé
	2SG	tu es allé(e)	tu as mangé
	3SG	il/elle est allé(e)	il/elle a mangé
	1PL	nous sommes allé(e)s	nous avons mangé
	2PL	vous êtes allé(e)s	vous avez mangé
	3PL	ils/elles sont allé(e)s	ils/elles ont mangé

Table 70. Present and perfect inflection of aller *'go' and* manger *'eat' in French*

In French, most verbs select *avoir* as their auxiliary, with the exception of about a dozen intransitive verbs (including *aller* 'go' and *venir* 'come') as well as all pronominal reflexive verbs, which select *être* (Bonami 2015:97). Perfective tenses enter into opposition with other tenses (here the present versus the past perfective). This is a clear instance of inflectional periphrasis.

Historically, such inflectional periphrasis comes from two different periphrases in late Latin. Both periphrases involved a verb *habere* 'have' or *essere* 'be' together with a past participle, with a resultative meaning (Ledgeway 2012; Squartini & Bertinetto 1995). The main periphrasis involved a verb of possession:

1) habeo paratam cenam
   I.have prepared dinner
   'I have prepared the dinner'

This construction came to acquire a resultative meaning in Latin in the present (Harris 1971; Price 1971:226; Bybee, Perkins & Pagliuca 1994). If I have the meal that has been prepared, then it is pragmatically inferable that I have prepared the meal. The Latin resultative construction was a present perfect, and could only be used with a participle of a transitive verb (Ledgeway 2012). The path of grammaticalization to a past perfect tense involved bleaching of the verb *habere* 'have',

which lost its lexical meaning of possession, and acquired a fixed coreference between the subject of the auxiliary and the subject of the lexical verb; progressively, the order of verb and past participle was fixed, and some of the agreement features of the past participle with the object of *habere* were lost (Squartini & Bertinetto 1995). The grammaticalization of this construction first occurred with verbs of cognition where it is impossible to interpret the experiencer subject of the participle and the possessor of *habere* as not having identical reference (Ledgeway 2012). In parallel, another periphrasis occurred with a copula and a past participle, with a similar result: *I am gone*, ended up meaning *I have gone*. This second construction tended to be found with intransitive verbs.

The two periphrases gave rise to various patterns of auxiliary selection: each lexeme selects one or the other auxiliary. This is the case in old Spanish, old Catalan, Occitan, French, Italian and most Italo-Romance dialects (Loporcaro 2016). Loporcaro (2007) indicates that this pattern of alternation was probably the common case in proto-Romance, which showed an active versus inactive alignment for perfect tenses: active verbs selected an auxiliary 'have', inactive verbs an auxiliary 'be'. The system was subsequently simplified in Spanish and Catalan, which only retain one auxiliary, 'have', for perfect tenses, and in Portuguese which ultimately selected auxiliary *ter* 'have' superseding the earlier auxiliary *haver*, as in Spanish.

The choice of auxiliary is thus correlated in a number of Romance varieties with the syntactic behaviour of the clause in which the lexical verb is used. Table 71 shows how a number of varieties select one or the other auxiliary depending on the alignment of the verb: the selection of auxiliary can be shown to operate in a hierarchical fashion across Romance varieties (Loporcaro 2007, 2016; in the tables, E=be, H=have). Syntactic alignment corresponds to the behaviour of a verb depending on the coding of its arguments. One can distinguish nominative-accusative alignment in which the subject of transitive and intransitive clauses pattern identically; absolutive-ergative alignment in which the subject of an intransitive verb and the object of a transitive verb pattern identically, singling out the subject of transitive clauses (ergative); and active-inactive alignment where the

subject of intransitives sometimes patterns with the subject of transitive verbs (unaccusative), and sometimes with the object of a transitive (unergative). Note that this alignment is lexeme based for the main lexical verb of the construction (unaccusative, reflexive, transitive, unergative). In some Romance languages, this lexical determination is shown further by the fact that some verbs which would be expected to behave in a similar way in fact select different auxiliaries. This is in particular the case in French, where *apparaître* 'appear' selects *être*, and *disparaître* 'disappear' selects *avoir*, despite their similar semantics (Bonami 2015:97). Auxiliary selection, although correlated with the lexical semantics of the verb, has to be recognized as partially arbitrary (Abeillé & Godard 2002; Bonami 2015), at least in French.

	INACTIVE					ACTIVE
	Unaccusative	Reflexive				Transitive / unergative
		retroherent	direct transitive	indirect unergative	indirect transitive	
Italian	E					H
Sardinian	E				H	
Picernese	E			H		
Old Florentine	E		H			
Engadine	E	H				
Spanish	H					

Table 71. Romance auxiliary options (Loporcaro 2016:814)

One is thus presented with systems which present a different morphological realization of the same set of feature values depending on the lexeme: that is, a system of inflectional classes which is highly correlated to some syntactic property of the lexeme (whether the verb is unaccusative, reflexive, unergative or transitive). Morphologically, systems of auxiliary selection in Romance languages can be considered as systems of inflectional classes which cross-cut the traditional system of classes based on thematic vowels and synthetic realizations. This fact is confirmed by the literature on the topic. Thus Bonami (2015) states that under his analysis of periphrasis, auxiliary selection is "reminiscent of the status of inflection classes: similar lexemes tend to cluster in the same classes, but there are exceptions. In the present approach, auxiliary selection is literally a matter of inflection class: just as different classes of lexemes may trigger the use of distinct rules of

synthetic exponence for the expression of the same feature, they may likewise trigger the use of distinct rules of periphrastic exponence" (Bonami 2015:97). Romance varieties with auxiliary selection may thus be treated as an instance of inflectional class through differential auxiliation, arising from the grammaticalization of two initially different structures in late Latin to yield a single grammaticalized periphrase for perfect tenses.

Such a system of inflectional classes can be compared to the concept of segregated inflectional class in Stump's theory. A segregated inflectional class is a system of inflectional class that only applies to a subset of the paradigm of a lexeme (Stump 2016). Stump takes as an example the inflection of verbs in Latin, which can be divided into three main stem formation rules. The inflectional classes concerning the *infectum* stem are the usual four classes distinguished by their thematic vowel. But these classes mostly do not align with the inflectional properties of class formation shown in the *perfectum*, which are mostly distinguished by the type of stem formation they exhibit[40]. Stump concludes that the best way to account for the data is to consider that this is a case of segregated inflectional classes, where each lexeme belongs to a given class for its *infectum* stem, and to another class for its *perfectum* stem. This amounts to considering that each of these classes is only a class of stem, not a class of lexemes. In the case of Romance auxiliary selection, there is no difference in stem formation: the choice of the auxiliary does not correlate with the choice of a given stem formation rule on the lexical verb (and completely crosscuts the rules of formation of the past participle in macro-classes). Up to a point, though, one could consider that the occurrence of periphrasis characterizes perfect tenses in those Romance varieties with auxiliary selection, and the occurrence of periphrasis could be formally distinguished with a new stem in a PFM model. In a different model, an inheritance hierarchy model for example, the first subdivision for verbs is their auxiliary selection, and then they are further subdivided into other inflectional classes, which has the advantage of preserving a conception of inflectional classes as classes of lexemes. This last model is the one adopted here.

40 This is the case for most classes except class 1 verbs which still exhibit a thematic vowel in the perfectum identical to that exhibited in the imperfectum.

Some varieties mix the two auxiliaries in the same paradigm, even though all verbs select the same paradigm for their perfective inflection. This is the case in a number of varieties of the Abruzzo region. In such varieties, there is no inflectional class distinction for perfect subparadigms, but a stronger formal distinction between the realizations attached to each feature value bundle. In some cells, the realization will be a 'have' auxiliary, in some a 'be' auxiliary. This is a fact about their morphology. Table 72 gives examples of the distribution of auxiliaries of such varieties. As one can see, the distribution of alternants is far from constituting natural classes. In such cases, these can be considered morphomic distribution of a suppletive type, although one would need diachronic data to prove it (Štichauer 2018).

	1SG	2SG	3SG	1PL	2PL	3PL
L'Aquila	E	E	H	E	E	H
Vasto	H	E	E/H	H	H	H
Introdacqua	H	E	H	H	H	H
Notaresco	E	H	H	H	H	H

Table 72. Mixed patterns in the Abruzzo (Loporcaro 2007:184)

In addition, there are varieties which exhibit heteroclisis, a type of mixed paradigm characteristic of systems of inflectional classes. These systems have been described by Loporcaro (2007) as triple auxiliation systems. They present for some verbs an auxiliary 'have', for others auxiliary 'be', and for yet other lexemes a mixed paradigm combining forms of 'have' and 'be'. Loporcaro (2007) describes the system displayed in the dialect of Telve Valsugana, spoken in the Trento province. In this dialect, unaccusatives select the auxiliary *essere*, and unergatives and transitives select the auxiliary *avere*. But the reflexives select a mixed pattern which shows heteroclisis between the *avere* type and the *essere* type. In effect, they display a different inflectional class. In this pattern, third persons realize perfect inflection selecting *avere* and the other persons selecting *essere*. Table 73 shows the relevant forms of the verb for a mixed pattern reflexive verb 'to get scared', where synchronically both auxiliaries are attested elsewhere in the system.

1SG	me son spaurá	E
2SG	te si spaurá	E
3SG	el s'a spaurá	H
1PL	ne son spaurái	E
2PL	ve sé spaurái	E
3PL	i s'a spaurái	H

Table 73. Heteroclite auxiliation in Telve Valsugana (after Loporcaro 2007:201)

Table 74 shows a summary of the system. Auxiliation is still linked to the lexical semantics of the verbs, but it shows a heteroclite paradigm for reflexives. A number of other Italian varieties show similar patterns (Loporcaro 2016:817-817).

INACTIVE		ACTIVE
unaccusative	reflexive	transitive / unergative
essere	heteroclite essere / avere	avere

Table 74. Heteroclite verbs in Telve Valsugana (Loporcaro 2007:202)

Systems of auxiliary selection in Romance can thus be considered as creating inflectional classes, albeit segregated inflectional classes[41]. Most systems show an opposition between two inflectional classes, each realized by a different auxiliary. Some systems in Italian dialects exhibit another distinction by presenting a heteroclite class. Such systems of inflectional classes arose from the grammaticalization and further convergence in meaning of two periphrastic structures in late Latin. The presence of heteroclite paradigms is a confirmation of their status as inflectional classes.

5.1.2 Basque auxiliated classes

The Basque verb is notoriously complex. For the sake of simplicity, I only analyse forms of the present indicative paradigm in this section, as it is enough to expose the difference between major inflectional classes. Similarly, I do not delve into the massive dialectal variation exhibited by verbs, but only present forms from the standard form of the language, called *Euskara Batua* (De Rijk 2008).

41 One can also consider them in a system of global inflectional classes, which only means that more classes will be distinguished. In an inheritance hierarchy model, the first division will be between perfective classes, and then each of these classes will subdivide for imperfective inflectional classes.

The system of Basque conjugations makes a sharp distinction between transitive and intransitive verbs (De Rijk 2008:116), each displaying different conjugations, in particular because their content paradigm is not identical: intransitives mark subject agreement in person, number, and, for close relationship second singular forms, gender of the addressee, while transitives mark agreement with both the subject and the object. The marking of subject and object follows an ergative-absolutive pattern: the object of transitive verbs is marked by the same set of affixes as the subject of intransitive verbs. In grammars of Basque, the verbal system is not traditionally analysed in terms of inflectional classes: I present a first analysis in those terms and I show that distinctions in the marking of plurality as well as the opposition between periphrastic conjugations and synthetic conjugations do create different inflectional classes in Basque, whose origins can be traced through internal reconstruction.

Cross-cutting the distinction in transitivity is another, purely formal, major distinction in the inflection of Basque verbs. A restricted number of verbs show synthetic conjugations in the present: all agreement markers as well as tense and aspect markers are realized within a single word form. Some of these verbs are transitive, others are intransitive. In addition, they do not all follow the same realization rules for their inflectional marking. Only about two dozen verbs inflect synthetically in present day Basque, but there used to be more in the earliest times of written attestations of Basque: Leizarraga's 1571 translation of the New Testament contains about 50 (De Rijk 2008:116). The synthetic verbs tend to be high frequency lexemes ('know', 'see', 'eat', 'say', etc.). In addition to these synthetic verbs, the majority of Basque verbs inflect with an auxiliary construction consisting of one synthetic verb, the auxiliary, which varies depending on the transitivity status of the lexical verb, and a non finite form of the lexical verb.

In this section, I first describe the synthetic conjugations, then the periphrastic conjugations, and analyse them as inflectional classes linked with the transitivity of the verb. In the second part, I trace the origins of such a system of inflectional classes.

In *Euskara Batua* there are six synthetic intransitive verbs, plus an additional four which also

require a dative argument (De Rijk 2008:116). There are also 18 synthetic transitive verbs.

VERB	joan	etorri	ibili	etzan	egon	izan
STEM	-oa-	-torr-	-bil-	-tza-	-go-	-iza-
MEANING	go	come	walk	lie down	stay	be / copula
1SG	noa	nator	nabil	natza	nago	naiz
2SG.CLOSE	hoa	hator	habil	hatza	hago	haiz
3SG	doa	dator	dabil	datza	dago	da
1PL	goaz	gatoz	gabiltza	gautza	gaude	gara
2SG.NEUTRAL	zoaz	zatoz	zabiltza	zautza	zaude	zara
2PL	zoazte	zatozte	zabiltzate	zautzate	zaudete	zarete
3PL	doaz	datoz	dabiltza	dautza	daude	dira

Table 75. Synthetic intransitive verbs in Basque (De Rijk 2008:120-123)

The synthetic conjugations of intransitive verbs are shown in Table 75. There are some evident similarities between the various forms, and between the verb form and the personal pronouns. Thus first person singular pronoun *ni* corresponds to a first person singular marking *n-* or *na-*. Then there is an evident relation between forms that were originally singular and forms that were originally plural. The second singular neutral form was originally a plural form which evolved first as a marker of politeness for referring to second person singular, much in the same way that English *you* evolved from an originally plural form. A second person plural form was reinnovated by adding a suffix *-te* to the second person singular form. All originally plural forms show a pluralizing element, which divides the verbs into different inflectional classes: *-z* for *joan* and *etorri*, *-tza* for *ibili*, an infixed *-u-* for *etzan* and *egon* (with an additional suffix *-de* for *egon*), and a suppletive base for *izan*. One can thus hypothesize that these synthetic forms were initially created through the grammaticalization of personal pronouns as agreement markers, plus some other elements indicating plurality (see Gomez & Sainz 1995).

1SG/3SG	daukat	1SG/3PL	dauzkat
2SG.CLOSE.NF/3SG	daukak	2SG.CLOSE.NF/3PL	dauzkak
2SG.CLOSE.F/3SG	daukan	2SG.CLOSE.F/3PL	dauzkan
3SG/3SG	dauka	3SG/3PL	dauzka
1PL/3SG	daukagu	1PL/3PL	dauzkagu
2SG/3SG	daukazu	2SG/3PL	dauzkazu
2PL/3SG	daukazue	2PL/3PL	dauzkazue
3PL/3SG	daukate	3PL/3PL	dauzkate
2SG.CLOSE.NF/1SG	naukak	2SG.CLOSE.NF/1PL	gauzkak
2SG.CLOSE.F/1SG	naukan	2SG.CLOSE.F/1PL	gauzkan
3SG/1SG	nauka	3SG/1PL	gauzka
2SG/1SG	naukazu	2SG/1PL	gauzkazu
2PL/1SG	naukazue	2PL/1PL	gauzkazue
3PL/1SG	naukate	3PL/1PL	gauzkate
1SG/2SG.CLOSE	haukat	1SG/2SG	zauzkat
2SG/2SG.CLOSE	hauka	2SG/2SG	zauzka
1PL/2SG.CLOSE	haukagu	1PL/2SG	zauzkagu
3PL/2SG.CLOSE	haukate	3PL/2SG	zauzkate
		1SG/2PL	zauzkatet
		2SG/2PL	zauzkate
		1PL/2PL	zauzkategu
		3PL/2PL	zauzkatete

Table 76. Present paradigm of the synthetic transitive verb eduki *'possess' (De Rijk 2008:195)*

The synthetic conjugation of transitive verbs is shown in Table 76 for the verb *eduki* 'have, possess', whose paradigm is relatively regular[42]. The verb marks plurality (for the object) through the use of an infix *-z-*, with a main stem *-uka-*, which gives a plural stem *-uzka-*. In *Euskara Batua* there are 18 transitive synthetic verbs (De Rijk 2008:187). The additional complexity is that transitive verbs agree with both their subject and their object. Thus the form *da-uka-t* 3SG.OBJ-have-1SG.SUBJ means 'I possess it'. Subject and object are always disjoint in Basque non-reflexive verbs, which explains the absence of some forms in this paradigm (there is a separate paradigm for reflexive verbs). In transitive verbs, prefixes mark the object, with forms similar to the subject of intransitive verbs. Subject agreement for transitive verbs is marked through a set of suffixes. In fact, prefixes mark the absolutive argument, while suffixes mark the ergative argument in an absolutive-

42 In the table, glossings indicate first the person, number and gender of the subject (suffix), and then the person and number of the object (prefix). The first column of forms presents forms with a singular (object) stem, the second column forms with a plural object stem.

ergative alignment. The set of ergative suffix agreement markers is shown in Table 77, following an analysis in terms of underlying forms by De Rijk (2008). One should note that there are a number of phonological rules adapting the actual surface form of the suffix. In particular, suffixes ending in *-a* lose it when word final, which triggers devoicing of its consonant (thus the first singular suffix surfaces as *-t*, except in the plural of **ion* 'say'). Thus **da-uka-da* surfaces as *da-uka-t*.

1SG	-da
2SG.CLOSE.F	-na
2SG.CLOSE.NF	-ga
3SG	-Ø
1PL	-gu
2SG.NEUTRAL	-zu
2PL	-zue
3PL	-te

Table 77. Subject suffixes on transitive synthetic verbs in Basque (De Rijk 2008:189)

Just like for intransitive synthetic verbs, transitive synthetic verbs fall into different inflectional classes depending on the form realizing (absolutive) pluralization on the stem. There are five different inflectional classes for synthetic transitive verbs, some of which only contain one lexeme. Thus **edun* 'have' forms its plural stem by prefixing *it-* to the singular stem *-u-*; *eduki* 'possess' with an infix *-z-* inside the singular stem *-uka-*, which gives a plural stem *-uzka-*; the verb **ion* 'say' realizes plural object through the addition of a suffix *-z* after the subject agreement marker; verbs of knowing and perceiving form their plural stem by suffixing *-zki* to the singular stem; and finally all remaining synthetic verbs suffix *-tza* to the singular stem to form their plural object stem (De Rijk 2008:190). For these inflectional classes, there is a small degree of semantic determination (verbs of knowing and perceiving), but otherwise no external constraint on the choice of inflectional class. These verb classes are not traditionally envisaged as inflectional classes in Basque grammars. But they are groupings of lexemes which inflect in the same way for identical feature values, and can thus be considered as inflectional classes. Note that the plural object stem is in fact now morphomic, as the originally second person plural has now become a second singular neutral form: it can no longer be characterized in terms of plural semantics.

Most verbs belong to periphrastic conjugations and do not present any synthetic forms. These

are characterized by the fact that they have a different auxiliary depending on the transitivity status of the verb: this is because different transitivity statuses imply a different content paradigm. Thus, in the present, intransitives use the intransitive auxiliary *izan* 'be', and transitives use the transitive auxiliary **edun* 'have'. Transitivity is by no means always lexically specified: the vast majority of Basque verbs allow for both intransitive and transitive use (De Rijk 2008:137). De Rijk analyses these verbs as basically intransitive, and allowing for a zero marked causative derivation. Periphrastic conjugations appear with a participle form of the lexical verb, which can be either perfect, imperfect or future. Auxiliated classes are thus also different from synthetic classes in that their paradigm distinguishes aspect, in addition to tense and mood. Thus for the intransitive verb *erori* 'fall', one can distinguish between present perfect (Example 2) and present imperfect (Example 3), a distinction which is not available to synthetic verbs. Synthetic and auxiliated verbs thus present a different content paradigm, which would in theory prevent us from considering them as two inflectional classes in the same system.

2) Ibaira erori da
river.LOC fall.PERF is
'He has fallen into the river'

3) Ibaira erortzen da
river.LOC fall.IMPERF is
'He is falling into the river'

The same aspectual distinction is available for transitive verbs, as shown in Examples 4 and 5:

4) Jonek etxea saltzen du
John.ERG house.ABS sell.IMPERF has
'John is selling the house'

5) Jonek etxea saldu du
John.ERG house.ABS sell.PERF has
'John has sold the house'

The system of Basque verb conjugations thus presents a number of different paradigms, which differ in their content paradigm. These can be treated as major distinctions of inflection. There is an opposition between transitive and intransitive paradigms, and in addition an opposition between synthetic and periphrastic paradigms. The distinction in content paradigm between synthetic and periphrastic paradigms is different from the distinction in transitivity in that it involves underdetermination rather than a different content paradigm. Synthetic verbs are underdetermined for aspect: no aspectual distinction is marked on them, which makes it possible to consider them inflectional classes of the same system. In addition, the synthetic paradigms also present different inflectional classes depending on their plural stem formation rules. It is possible to formalize the system of Basque inflectional classes as in Figure 1, which represents the inheritance hierarchy for Basque verbs.[43]

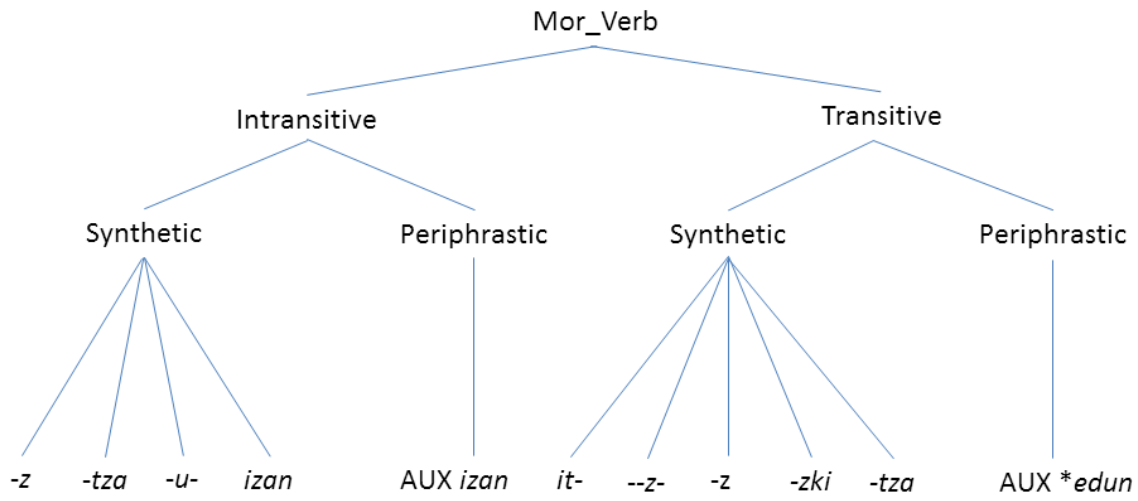


Figure 2. The morphological network of Basque inflectional classes

The system of conjugation that is seen in contemporary *Euskara Batua* was already in place by

43 I leave out the marking of dative arguments from this analysis for the sake of simplicity. This would add a further division in terms of transitivity, for verbs that are intransitive but also take a dative argument, and ditransitive verbs. Thus ditransitives for example make use of the auxiliary **edun* 'have' in a paradigm that marks absolutive, ergative and dative arguments, as in the following example:

Mutil	honek	neskatxa	horri	egun	zoriontsua	opa	dio
boy	this	girl	that	day	happy	wish.PERF	AUX

'This boy wishes that girl a happy day' (De Rijk 2008:343)

the time the first written records of Basque appeared in the sixteenth century. It is thus not possible to use historical records to reconstruct the origins of the system. The only valid piece of data that we have is the fact that there used to be more synthetic verbs in the sixteenth century than there are now: the numbers have been reduced by half. This may be an indication that synthetic verbs are older than the analytic, auxiliary constructions. There are two main theories in trying to reconstruct the proto Basque verb system internally, using the knowledge we have about the system and knowledge about typological change in verb systems.

The first theory is that in proto Basque there were only synthetic verbs, and that the auxiliated class grammaticalized later (Trask 1997; Gómez & Sainz 1995). A number of arguments can be adduced in favour of this theory, notably the fact that synthetic verbs present non-derived participles, whereas the vast majority of verbs in the auxiliated class are derived with a perfect participle marker (Trask 1995). In favour of that theory is also the fact that more verbs used to be conjugated synthetically in the early literary period. Gómez & Sainz (1995) indicate that contact with Latin and the Romance languages (Aragonese, Spanish, Gascon) "must have at least promoted and accelerated the rapid development and expansion of the system of periphrastic forms, if it did not in fact give rise to it" (Gómez & Sainz 1995:238). Padilla-Moyano (2013:354-356) comes to the conclusion that the Romance model was not decisive for the rise of all periphrasis in Basque, on the grounds that only the perfective periphrasis offers a valid model to a similar structure. He concludes that "such an influence could have prompted the emergence of parallel structures in Basque, or simply the increase in frequency of a pre-existing periphrasis" (Padilla-Moyano 2013:356). It may thus be possible that the analytic forms arose through contact with Romance varieties, at least for the indicative periphrases in participle plus auxiliary (see Mounole Hiriart-Urruty 2006), a fact that would posit that synthetic forms are more ancient than the analytic construction. In any case, it is highly probable that at the time of this contact, the synthetic conjugation was no longer productive, in light of the large number of verbs borrowed from Latin in the periphrastic conjugation (forms of which gave rise to perfect participles in *-tu* borrowed from Latin participles in *-tum*). Synthetic

verbs grammaticalized from a verbal complex consisting of a verbal root and various pronominal and tense elements, which progressively cliticized (Gómez & Sainz 1995:250). Thus, a present marker **da=* 'now' was first added as a present tense marker, in addition to subject and object pronouns which cliticized to the root, giving the present day agreement markers that bear a close resemblance to pronouns. As Gómez & Sainz (1995:250) put it, "free pronouns with the same phonological shape that they present nowadays (*ni*, *hi*, *gu*, *zu*) were incorporated to the verb via cliticization, and (...) all the differences between free pronouns and agreement markers are due to phonological change: loss of intervocalic *-d-* and monophthongization" (see also Ariztimuño 2013:365-366). This is again assuming, as I have done in Chapter 4 about the grammaticalization of bound pronominals in South Halmahera-West New Guinea languages, that grammaticalization processes involving cliticization and affixation always involve some degree of sound change.

But that theory only gives information about the relative time frame of the two constructions, synthetic and periphrastic, not about their origins. The second theory has been proposed more recently by Joseba Lakarra (2006; also see Ariztimuño 2013). It states that synthetic verbs first arose in proto-Basque from the grammaticalization of asymmetrical serial verb constructions, constructions where more than one verb is present, the first being a minor verb expressing aspectual distinctions ('sit' for continuity, 'finish' for completeness, etc.), the second being a major verb bearing the main lexical aspect. We would thus have constructions such as **dar tor* sit-come which would come to mean 'to be on coming'. The two verbs are assumed to have grammaticalized, the first giving a temporal affix (progressive aspect to imperfect to present) *da-* which is found nowadays in the third person present of synthetic verbs (Ariztimuño 2013:424-425). For first and second person, a pronoun was added to the construction and first cliticized then affixed as an agreement marker. The plural markers were added later, distinguishing multiple inflectional classes. In such a scenario, the ancestor language of Basque was isolating, and acquired synthetic forms of the verbs through the grammaticalization of serial verb constructions. The synthetic forms of the verbs ceased at one point to be productive, and some synthetic verbs grammaticalized to become

auxiliaries in periphrastic constructions (Ariztimuño 2013:427).

The two theories are not completely irreconcilable. They have in common that processes of grammaticalization ended up creating different inflectional classes in the language They differ on points of detail, such as the origins of third person present marker *da-* which in one theory is ascribed to an adverb meaning 'now', and in the other is given a verbal origin in a serial verb construction. But such differences do not bear any significance for the creation of inflectional classes. More importantly, both theories state that synthetic verbs grammaticalized before periphrastic verbs, although the serial verb construction theory does not assume that all verbs grammaticalized a synthetic conjugation before the periphrastic conjugation grammaticalized. In all probability, it is possible to unify the diachronic scenario given by both theories. First synthetic verbs grammaticalized the present marker for third person, and agreement markers for first and second persons. At that stage, there may have been an opposition between inflecting verbs and non-inflecting verbs, with all inflecting (synthetic) verbs following identical realizations. These verbs may have entered into complex serial verb constructions where only one of the verbs inflected for absolutive and ergative agreement. Later, plural absolutive elements were added, which created different inflectional classes depending on the pluralizer chosen for each verb. Finally, at a later stage, probably starting from a serial verb construction in which only one of the verbs was inflected, a periphrastic construction grammaticalized with a different auxiliary for intransitive and transitive verbs, giving rise to the periphrastic inflectional classes. This last development may have been triggered or at least favoured by contact with neighbouring Romance varieties presenting a periphrastic perfect.

5.2. Cycles of bound pronominals grammaticalization

In addition to auxiliation, a language can acquire inflectional classes through the grammaticalization of pronominals as subject and object agreement markers. In this section, I first

provide an account of the grammaticalization of pronouns into agreement markers on the verb in general, before examining how the grammaticalization of such forms gave rise to inflectional classes in Skou and in Arapesh, non Austronesian languages of New Guinea, which show an interesting cycle of grammaticalization that has not until now been described as a potential source of inflectional classes.

5.3.1 The grammaticalization of bound pronominal forms

The grammaticalization of pronouns as agreement markers on the verb has already been alluded to in this thesis in Chapter 4 about South Halmahera-West New Guinea languages, as well as in the preceding section about Basque. This is because grammaticalization processes giving rise to affixal marking are often, if not always, accompanied by some degree of sound change in the form of phonological reduction. The grammaticalization of pronouns as agreement affixes on verbs is attested in a wide range of languages.

The grammaticalization of pronouns as agreement markers shows the expected four parameters of grammaticalization, extension, desemanticization, decategorialization and erosion (Heine & Kuteva 2007). In addition, there is a cline of grammaticalization from fully independent pronouns to inflectional affix, through various stages of cliticization. It shows extension in that in a first stage, pronouns are used with a wider range of constructions, mainly appearing together with a full subject NP. Decategorialization then means that progressively pronouns lose their capacity for reference, and become more grammatical as purely agreement markers, which also concerns desemanticization, or semantic bleaching. Furthermore, for decategorialization, these pronominal forms lose the possibility of appearing non-adjacent to the verb complex: their ordering possibilities become limited to specific positions immediately adjacent to the verb, in a process of cliticization that is completed when the pronominals lose their phonological independence, in particular their ability to bear stress. Finally, some phonetic erosion occurs, which changes the shape of the pronominal marker depending on the phonological context of attachment.

In the context of the creation of inflectional classes, the most important element in this process

of grammaticalization is phonetic erosion, because it can create phonetic alternants depending on the phonological profile of the verbal root it attaches to. Thus, in South Halmahera-West New Guinea languages, the shape of the marker became different when the verbal root was consonant initial or vowel initial. The process can create separate inflectional classes highly correlated with phonology. This is by no means an automatic feature: Basque, for example, grammaticalized agreement markers on finite verbs without it giving rise to allomorphy (see 4.1.2.). The languages studied here show a different pattern in the form of a cycle of two such grammaticalizations of bound pronouns, which creates inflectional classes.

5.3.2 Skou

Skou is a Macro-Skou language spoken in the north of Indonesian Papua (Donohue 2002, 2003). Its verbal morphology shows interesting patterns of allomorphy for subject agreement, which are analysed in a diachronic perspective by Donohue (2003). Donohue's analysis shows in particular that the patterns of allomorphy can be explained by two successive cycles of grammaticalization of pronouns first into clitic bound pronominals, then into affixes marking pronominal agreement.

Skou verbs agree with their subject in person, number, and gender. There is more than one pattern of allomorphy for the subject prefixes. Agreement markers can be the sole marker of the subject in the absence of a full subject NP only when the subject is topical (Donohue 2003:479-480). Skou also presents the peculiarity of having a strict (C)V syllable structure, a fact which is relevant to the analysis given by Donohue. In addition, Skou is a tonal language.

All verbs in Skou inflect for subject agreement by means of a proclitic. Approximately a third of verbs in the language only show proclitics as agreement (Donohue 2003:481). These proclitics can be shown to be different from full pronouns in that they do not distinguish dual number: the plural proclitic is used instead. Table 78 shows the paradigm of inflection for the verb *e* 'board, get onto, travel by means of':

1SG	nì=e
2SG	mè=e
33SG.NF	ke=e
3SG.F	pe=e
1PL	ne=e
2PL	e=e
3PL	te=e

Table 78. The verb e *'board' in Skou (Donohue 2003:481)*

In addition to the proclitic marking, a number of verbs mark subject agreement through a set of prefixes which also corefer with the subject person, number and (in the third person singular), gender. These verbs thus mark subject agreement twice, through a clitic and a prefix:

6) ápolè-ha ne=n-ang ka

tulip-leaf 1PL=1PL-eat NEG

'We can't eat tulip leaves' (Donohue 2003:480)

The underlying forms of the prefixes are shown in Table 79 (Donohue 2003:482). These forms are very close to the equivalent consonant in the free pronouns and the proclitics. The underlying forms are also those which realize subject agreement through prefixes in vowel initial verbs (see Table 80 below).

	SG	PL
1	Ø-, k-, n-	n-
2	m-	Ø-,
3.NF	k-	t-, y-
3.F	p-	

Table 79. Underlying forms of subject agreement prefixes in Skou (Donohue 2003:482)

The variation shown in prefixes for first person singular only concerns three vowel initial verbs. Consonant initial verbs have no prefix for that cell. Donohue argues that the underlying forms of the prefixes clash with consonant initial roots and the strict CV syllable structure of the language, yielding reduced forms of the consonant clusters, which create the inflectional classes shown in Table 180 (the forms are shown without the additional proclitics, which are identical for all verbs). These classes correlate to a large extent with phonological shape, depending on the first segment of the root (as shown in the first singular form). The classes also show extensive syncretism.

	vocalic	bilabial	alveolar-*l*	alveolar-*r*	velar	glottal
	'go east'	'get.F'	'release'	'go'	'get'	'walk'
1SG	e	wé	lú	re	ké	ha
2SG	me	pé	pú	me	bé	ma
33SG.NF	ke	wé	lú	ti	ké	ka
3SG.F	pe	wé	rú	te	wé	wa
1PL	ne	wé	rú	ne	ké	na
2PL	e	wé	lú	re	ké	ha
3PL	te	wé	rú	te	ké	ya

Table 80. Prefixal inflectional classes in Skou (Donohue 2003:483)

Donohue's analysis of these data is as follows. In Skou languages, there was a first grammaticalization of subject pronouns as agreement markers which created prefixes. This is confirmed by similar developments in other Macro-Skou languages. A peculiarity of Skou was the drastic reduction of these markers following the reduction of initial consonant clusters. About a third of verbs ended up without any realization of subject agreement in their inflected forms. This favoured a second wave of grammaticalization of pronouns as agreement markers, which created proclitics on all verbs. This would explain why some verbs realize subject agreement through both proclitics and prefixes.

The situation in Skou is thus the creation of inflectional classes following a cycle of two grammaticalizations of bound pronominal markers. What is peculiar in that situation with respect to other known grammaticalization patterns of subject agreement markers in Papuan or Austronesian languages, is the fact that there were two waves of a similar process, which created an additional class with only proclitics, not prefixes.

5.3.3 Arapesh

Arapesh is a language of the Arapesh family belonging to the Torricelli phylum, spoken in the East Sepik province of Papua New Guinea (Hammarström et al. 2017). Arapesh is better known for its gender system and its system of noun classes on nouns (see Chapter 8), but I show here that it is also remarkable for a very rare type of inflectional class on its verb system: the distinction of inflectional classes through the position class of affixes, a type of class which has not been commented upon. This section is based on the data given in the description of the language by

Fortune (1942).

Arapesh marks subject agreement on all verbs through prefixes, for both transitive and intransitive verbs. For transitive verbs, there are two different classes depending on where the object agreement is realized: one class marks object agreement through prefixes (transitive class 1), the other through suffixes (transitive class 2) (Fortune 1942:45ff.). It is not possible from the data given by Fortune (1942) to find any further determinant for the membership of verbs in each class: the choice of alternative realizations of object marking does not seem to be constrained by syntax or phonology. First and second person object agreement for both classes is given in Table 81.

	PRO	Class 1	Class 2
1SG	eik	ja-	-ei'
2SG	ɲak	ɲa-	-eɲ
1DU	awhok	who-	-ouh
1PL	apak	ma-	-ap
2PL	ipak	pa-	-ip, -eip

Table 81. Inflectional classes in Arapesh for first and second person objects (Fortune 1942:45)

For third person object, the system presents differential object marking based of the feature of humanness: human objects distinguish between masculine and feminine (Table 82), and non-human objects mark gender along one of thirteen different classes for each inflectional class (Table 83). For these markers, the link between the prefix variant and the suffix variant is quite evident: the markers are nearly identical (and similar to the noun class markers), except that prefixes end in a vowel *-a* while suffixes start with that vowel: they are thus in mirror ordering. They are also quite similar to the noun class marker on nouns.

	Class 1	Class 2
3SG.MASC	na-	-an
3SG.FEM	kwa-	-oku̥
3PL.MASC	ma-	-um̥
3PL.FEM	wa-	-u, -ou
Mixed company[44]	ʃa-	-eiʃ

Table 82. Arapesh inflectional classes for human objects (Fortune 1942:46)

[44] The forms used for a mixed company of humans are identical to the gender VIII of inanimates.

	Noun class markers		Prefixal object		Suffixal object	
	SG	PL	SG	PL	SG	PL
I	-b	-bys	ba-	bysa-	-ab	-abys
II	-bør	-røb, -ryb	børa-	røba-	-bør	-røb
III	-g	-gas, -as	ga-	gasa-	-ag	-agas
IV	-kʊ̥	various, -u	kwa-	wa-	-okʊ̥	-ou
V	-m	-eip, -ip	ma-	pa-	-am	-eip, -ip
VI	-n	-b	na-	ba-	-an	-ab
VII	-n	-m	na-	ma-	-an	-um
VIII	-ɲ	-ʃ	ɲa-	ʃa-	-eiɲ	-eiʃ
IX	-p	-s	pa-	sa-	-ap	-as
X	-r, -l	-guh̥	ra-	guha-	-ar	-aguh
XI	-t	-gʊ̥	ta-	gwa-	-at	-agʊ̥
XII	-ʊ̥h̥, -uh	-ruh	wha-	ruha-	-oʊ̥h̥	-aruh
XIII	-ah̥ -uh̥	-eh̥ -ih̥	ha-	ha-	-ah	-eh

Table 83. Inflectional classes on Arapesh verbs for non-human objects (Fortune 1942:47)

How does such a system come into existence? The first thing to note is that markers are quite similar in both classes, other than their difference in position. They can also be related to other markers, in particular pronouns or class markers. Thus for first and second person objects, the affixal markers bear a close relationship to the full (intensive) form of the corresponding pronoun, as shown in Table 83. One can easily imagine that these pronominal agreement markers arose from the grammaticalization of full pronouns in object position. There must have been some discourse or syntactic factors for object pronouns to appear before or after the verb, for example whether the object was topicalized. In any case, there was a differential grammaticalization of such pronouns, some emerging as prefixes, other as suffixes, depending on the transitive verb chosen. The result is a system where inflectional classes are distinguished by the position class of the affixes, a situation that to my knowledge is extremely rare in the languages of the world[45]. It seems to be a feature of Arapesh languages, as it also seems to appear in Bukiyip (Conrad & Wogigia 1991:14-15).

45 One other possible example I can think of is Bitur, a Papuan language from south New Guinea, which seems to distinguish two classes of transitives by the position of its object affix in the template. Philip Rogers, who is preparing a PhD involving fieldwork on that language tells me that the system is very complex and that it is premature, seeing the state of our knowledge, to confirm that this is indeed the case (Philip Rogers p.c.). Such classes are also attested in Qafar.

5.3. Conclusion

This chapter has confirmed that inflectional classes can arise through grammaticalization. Cases already known from the literature concerned the further grammaticalization of existing affixes into inflectional class markers, and the grammaticalization of subject and object pronouns into agreement markers as possible sources. In addition to these cases, I have shown that the grammaticalization of an auxiliary construction can also give rise to inflectional classes, either by distinguishing two different auxiliaries for different verbs (Romance and Basque), or by opposing a synthetic conjugation to a periphrastic conjugation of specific lexemes (Basque). Non-Austronesian languages of Papua further reveal that there can be very specific cases of the grammaticalization of bound pronominals giving rise to inflectional class distinctions. This can either come in the form of two successive waves of grammaticalization of bound pronominals, leaving some verbs with one marker, others with two markers (Skou), or in the differential grammaticalization of object pronominals in different syntactic positions depending on the verbs, which creates a rare form of inflectional class by affix position. I now turn to the third origin of inflectional classes mentioned in the literature, reanalysis.

6. Reanalysis

This chapter examines cases of reanalysis which create inflectional class distinctions. It expands on the already attested cases by providing data from a wider set of languages.

Reanalysis in morphology corresponds to a process of change which is "a new way in which speakers understand the structure of a word by relating it to other words in a different, novel way" (Haspelmath 1995). This type of change thus affects the way in which speakers analyse the internal structure of words. It is a change in their underlying structure which does not affect (or does not necessarily affect) the surface realization of words. As such, it has to do with the problem of word segmentation. In morphological reanalysis, a segment or a group of segments that were originally part of the stem come to be reanalysed as being part of an affix. This change in the way speakers segment the word is termed secretion by Haspelmath (1995). Secretion has been known in the literature to create inflectional classes in some Germanic varieties (Dammel 2011). In this chapter, I present more examples of reanalysis of the secretion type from a broader range of languages including Austronesian (Maori, Manam), Pama-Nyungan, and Romance (Romanian).

6.1. Secretion

Secretion is the the subtype of reanalysis "where a non-affixal part of a root is reanalyzed as part of an affix" (Haspelmath 1995:8). It corresponds to what Dammel (2011) calls reanalysis. It very often includes some sound change to the base form of a lexeme, which leaves some additional phonological material in derived forms. Such changes often display morphotactic opacity, which according to Haspelmath (1995) is a major trigger for reanalysis, together with affix syllabicity. As Haspelmath puts it, "in such cases where the base form is reduced phonologically, it is usually reinterpreted by speakers as the underlying stem form, and the affix which preserves the original stem consonant is reanalyzed as being part of the affix" (Haspelmath 1995:16). Such reanalysis of

part of the original stem material as affixal material is at play in the creation of inflectional classes in Maori, in Manam, in Pama-Nyungan languages and in Romanian. With these examples, I show that reanalysis is a major source for inflectional classes, and that in particular it tends to create thematic elements as class identifiers, as is the case in Maori and in Manam.

6.2. Maori

Maori is the textbook example of reanalysis giving rise to inflectional classes. The historical development of reanalysis in Maori also involves sound change, which is the reason why Carstairs-McCarthy (2010) chooses this example as a case of inflectional class arising from sound change, when properly it involves first a sound change creating morphotactic opacity, then reanalysis (see Chapter 4). It has been the subject of much debate between the proponents of a phonological solution (Sanders 1990; Sanders 1991; Parker Jones 2008) and proponents of a morphological solution involving multiple inflectional classes (Hale 1973; Hale 1991; Blevins 1994; Blevins 2008). In this section, I review the evidence, and opt for a morphological analysis whereby reanalysis historically produced a system of inflectional classes.

The data are simple. In Maori, there is a consonant versus zero alternation between the active form of the verb (no consonant) and the passive and gerundive forms of the verb which exhibit a thematic consonant. Table 84 shows in its first column the formal relationship between active and passive, then the active and passive form of each type of verb in Maori. Twelve different types can be can be distinguished depending on the presence or absence of a thematic consonant (*maka~makaia, whiu~whiua* without thematic consonant) and the type of consonant involved.

Construction	Active	Passive	Gloss
X:Xia	maka	makaia	throw
X:Xa	whiu	whiua	chase
X:Xtia	awhi	awhitia	embrace
X:Xkia	hopu	hopukia	catch
X:Xmia	inu	inumia	drink
X:Xina	aroha	arohaina	love
X:Xna	tahu	tahuna	burn
X:Xngia	tohu	tohungia	point out
X:Xnga	kai	kainga	eat
X:Xria	mau	mauria	carry
X:Xhia	kimi	kimihia	seek
X:Xwhia	whao	whaowhia	put into

Table 84. Active and passive verbs in Maori (Sanders 1990:151)

Hale (1973) was the first to distinguish two possible analyses for the presence or absence of thematic consonants in Maori verb forms. He distinguished a phonological analysis stipulating that roots have an underlying final consonant which is phonetically deleted in word final position (thus in the active form), from a conjugational analysis stating that the base form of the verb is the active form, and that Maori presents twelve inflectional classes showing allomorphy of the realization of the passive marker, each class being characterized by a different thematic consonant. He clearly favoured the second option, for reasons having to do with the phonotactics of Maori (strict CV syllable elsewhere), and the productivity and default realization of the passive suffix *-tia*, which is extended to all contexts where a speaker does not remember the class a given verb belongs to, and to loanwords. All derived causatives, and nouns converted into verbs also take the suffix *-tia* in the passive. Hale concludes that "these facts are entirely consistent with the suggested reanalysis" (Hale 1973:417).

Historically, the pattern of consonant alternation in Maori originated in sound change. In particular, final consonants were deleted earlier in the history of Maori. This means that in the active form only the final consonants were deleted. In the passive, consonants were protected by the presence of a suffix, and stayed. Speakers then applied reanalysis to these forms. The active form was felt to be the basic form of the lexeme, and the now thematic consonant was ascribed to affixal material, creating a number of inflectional classes. Then one of the suffixes, probably because of its

frequency, was chosen as a default affix *-tia* used for new lexemes which did not originally present a thematic consonant (derived forms and loanwords). One interesting point arises with this type of inflectional class creation. In Chapter 4, I have shown that sound change had a tendency to split existing inflectional classes into two new classes, as was the case for the Francoprovençal reflexes of Latin first conjugation. In the Maori case of reanalysis on the contrary, one witnesses the simultaneous creation of twelve different inflectional classes. Reanalysis thus creates classes with a different form than simple sound change.

The same type of consonant alternation with a similar historical origin is found in a number of other Oceanic languages, as is shown in the next section for Manam.

6.3. Manam

A case similar to the classical Maori example but with a more complex system of classes is exhibited by Manam, an Oceanic language spoken in the Solomon Islands (Lichtenberk 1983). Manam verbs mark a number of inflectional features, including subject and object agreement. There are two main paradigms of Manam verbs with a different content paradigm, intransitive and transitive verbs. Manam presents allomorphy at two levels. First there is allomorphy of the object marker, restricted to third person plural forms. Then there is the presence in a number of forms of a thematic element which creates an extended stem through the addition of a consonant.

Just like Maori, Manam presents a system of thematic consonants, which appear when the stem is followed by a suffix. Thus in Example 1, no thematic consonant appears, while in Example 2, because there is a suffix, it appears:

1) i-nóʔu

   3SG-jump

   'He jumped'

2) ári i-nóʔu-l-i

fence 3SG.SUBJ-jump-THC-3SG.OBJ

'He jumped over the fence'

Not all verbs present a thematic consonant: a sizable number of verb stems end in a vowel. The possible thematic consonants only consist of a subset of the consonant inventory of Manam: only /t, ʔ, s, m, n, ŋ, l, r/ appear, while /p, b, d, g, z/ never appear (Lichtenberk 1983:156-157). Lichtenberk gives an analysis of thematic consonants in Manam in which he refutes both a phonological analysis and a conjugation analysis. The phonological analysis consists in positing that thematic consonants are part of the stem, and that they are deleted in environments defined by the absence of a suffix (a morphologically conditioned sound alternation), or to define it phonologically, when they are word final. He refutes it on the grounds that Manam has a rather strict CV syllable, but that nasals can appear in coda, while they never appear as a thematic consonant in coda. There is also evidence from reduplication that thematic consonants are not part of the root of the verb (Lichtenberk 1983:144-145). What Lichtenberk calls the conjugation analysis posits that thematic vowels are somehow part of the affix, and that the allomorphy of all the affixes following the thematic consonant designates inflectional classes, i.e. that object suffixes and nominalizing suffixes show allomorphy. He refutes it on the grounds that short and long transitives (those showing the transitivizing suffix *-aʔ*) show the same thematic consonant: thus derivational material can occur between the thematic consonant and the inflectional suffix (Lichtenberk 1983:152-153). Lichtenberk finally analyses thematic consonants as a morpheme of their own whose sole function is to indicate the class of the verb. In modern terms, one can say that this thematic element serves to create new stems whose repartition is morphomic: it appears when the verb is inflected for object, when it is nominalized, and in long transitives. The thematic consonants form inflectional classes through the presence not of specific allomorphy from the inflectional material, but through their sole presence, much in the same way that thematic vowels act in Romance languages and in Latin. As in Latin, they are partially correlated to the allomorphy of inflectional suffixes (see below).

Manam also presents allomorphy for the third person plural object marker. Manam distinguishes for each verb between higher animate object (including humans) and other objects only in the third person plural cell. This is a type of marking that is similar to many cases of differential object marking attested in Oceanic languages. The marker for higher animals is *-di*, which is identical to the marker of non higher animals of class 3. Markers seem to be related phonologically. Class one verbs mark non higher animate third person plural objects by either *-i* or a zero affix (optionally zero after a nasal), class two by a zero affix, and class three by *-di* (Lichtenberk 1983:122). Note that for class one verbs, the realization of the third person plural object marker is syncretic with the third person singular marker present for all classes. The marker *-di* can be said to be the default realization of the third person plural object marker, as it is used for both higher animates in all classes, and for non higher animates of class three, which is also the productive class where loanwords get assigned.

There is a relationship between the allomorphy exhibited by the third person plural object marker and thematic consonants. In particular, for that cell, verbs of class two never mark a thematic consonant (Lichtenberk 1983:129), a fact that could be interpreted phonologically as the deletion of a final consonant in the absence of an affix protecting it (Example 4). In any case, the thematic consonant appears in other places in the paradigm of class two verbs (Example 3):

3) níu u-ʔózo-m-i
   coco 1SG-husk-THC-3SG
   'I husked the coconut' (Lichtenberk 1983:129)

4) níu u-ʔózo-Ø
   coco 1SG-husk-3PL
   'I husked the coconuts' (Lichtenberk 1983:129)

In addition, verbs which do not present a thematic consonant tend to belong to class three and mark 3PL object by *-di*; verbs containing the transitivity suffix *-aʔ* belong to class one, even when the suffix applies to an already transitive base; finally loan words belong to class three even when

they end in a consonant (Lichtenberk 1983:131).

The thematic consonants of Manam originate in root final consonants of proto-Oceanic which have been preserved inside words and lost word finally. This can be shown by comparing the thematic consonant of Manam verbs to the structure of verbs in the proto-language, and in particular their final consonant (Table 85).

Manam		proto-Oceanic	
toto-ʔ	'cut'	*totok	'hack'
ʔaʔa-s	'smooth'	*kakas	'scratch'
ono-t	'shut, close'	*ponot	'shut'
saʔi-l	'stamp'	*sakil	'stamp'

Table 85. Manam thematic consonants and proto-Oceanic roots (Lichtenberk 1983:147)

This also explains why there is a class of verbs (belonging to the third class of plural object allomorphy) which do not exhibit a thematic consonant. The athematic conjugation simply originates from verb roots in the proto-language which were vowel-final. The fact that some consonants never appear as thematic consonants reflects the fact that some consonant clusters, which gave rise to them, never appeared in coda position in the proto-language (Lichtenberk 1983:157).

The root final consonant was lost in some more simple environments in the inflection of Manam, in particular those contexts where there is no object agreement. This creates morphotactic opacity (Haspelmath 1995). Thus verb roots are reanalysed as being vowel-final based on simpler forms of the verb, and the thematic element is reanalysed as a specific formative indicating class. This is a clear case of secretion, where part of the root segments were reanalysed as belonging to the affixal domain. Some thematic consonants are then used to extend the verb lexicon, with some semantic generalizations. Thus all Manam transitive verbs based on a kinship term take *-m-* as a thematic consonant, in a pattern that is a clear innovation with regard to proto-Oceanic: *tama-m-i* 'regard someone as one's father' (Blevins 2008:96).

A situation similar to Manam and Maori is found in a number of Oceanic languages which present thematic consonants in parts of their verbal paradigm, as in Toqabaqita, Samoan, Tongan,

Niuean and Hawaiian (Blevins 2008:96).

6.4. Pama-Nyungan conjugations

Most Pama-Nyungan languages present a system of verbal inflectional classes. These classes are often characterized by a thematic element, a consonant, appearing between the root of the verb and the inflectional affixes in at least some of the forms of the verb. Thus, Yidiny presents three inflectional classes, characterized by the presence of a thematic consonant between the root and the inflectional suffix, as shown in Table 86.

	Class 1	Class 2	Class 3
Imperative	-n	-Ø	-rr
Present-future	-ŋ	-l	-r
Past	-nyu	-l-nyu	-r-nyu
Purposive	-na	-l-na	-r-na
Apprehensional	-n-ji	-l-ji	-r-ji

Table 86. Inflectional classes in Yidiny (Dixon 1980:383)

In Yidiny, Class 1 verbs are characterized by the presence of a consonant *-n-*, as well as the fact that it is an open class. Class 2 is also open, and characterized by a segment *-l-* occurring between the root and the inflectional affix. Class three is a closed class of about 20 members, characterized by a thematic consonant *-r-*.

Not all languages show a clear presence of a thematic consonant for all forms, as is the case in Yidiny, due to various sound changes having affected the root-affix boundary. Pama-Nyungan languages which show inflectional classes on verbs show between two and seven classes. Thus in Nyawaygi there are seven classes (Dixon 1980), which are distinguished by their pattern of allomorphy. Not all forms show a clear instance of thematic consonant, due to sound change, as shown in Table 87. The conjugations show a number of default forms: *-nya* for all classes in the unmarked positive; *-ga* for three classes in the imperative; *-wanya* for four classes in the recent past; *-gu* for two classes in the purposive. In addition, class 1b and 4 are open classes. It is interesting to note that class 1b also includes four default forms out of five forms. It can thus be considered as the

default class of verbs in Nyawaygi.

	1a (y)	1b (Ø)	2a (ŋ)	2b (m)	3a (n)	3b (n/ŋ)	4 (l)
Unmarked positive	-nya	-nya	-nya	-nya	-nya	-nya	-nya
Imperative	-yga	-ga	-ga	-ma	-na	-ga	-Ø
Irrealis	-yma	-ma	-jima	-njima	-njima	-jima	-lma
Recent past	-yanya	-wanya	-wanya	-wanya	-nmanya	-wanya	-lanya
Purposive	-ygu	-gu	-gagu	-magu	-nyagu	-nyagu	-gu

Table 87. Inflectional classes in Nyawaygi verbs (Dixon 1980:397)

In some languages, some classes have very restricted membership, to the point that some authors analyse them as consisting of irregular verbs. This is the case in Gumbaynggirr (Eades 1979). Table 5 presents the inflectional patterns of Gumbaynggirr; the regular pattern follows the model of the verb *panaa* 'arrive', and is the productive class; the other model verbs are in fact irregulars, that is classes with only one member.

	arrive	see	hit	take
Past	paana-ng	nyaa-wang	pu-waang	maa-ning
Present	paana-y	nya-yaki / nyaa-kay	pu-waang	maa-nyji
Purposive-Future	paana-y-ku	nyaa-ku / nyaa-kay-ku	puu-m-ku / pu-ma-yku	maa-n(i)-ku / maa-na-yku
Imperative	paana / paana-la	nyaa-(ya)ka / nya-yaki	pu-ma	maa-na / maa-ni
Future	paana-w	/	/	maa-ni-w

Table 88. Inflectional classes in Gumbaynggirr (Eades 1979:301)

There has been some debate about the origins of those inflectional classes. For Dixon (1980), they arose because of blurring of root affix boundaries, which in other words is reanalysis of the secretion type. On the contrary, for Harvey (2008), those inflectional classes arose through the further grammaticalization of previous TAM markers: inner affixes were further grammaticalized, semantically bleached, and reanalysed as part of the outer suffixes. In fact, the two main hypotheses proposed in the literature both account for the creation of inflectional classes in Australian languages through reanalysis. The question is whether it is of the secretion type (Dixon), or of the conglutination type (Harvey). Neither of these authors draws the distinctions I draw here for types of reanalysis based on Haspelmath (1995). In this section, I sum up the arguments for both sides of

the debate, and try to draw some conclusions.

Dixon (1980) posits a very simple structure for the proto-Australian verb consisting just of a root plus an inflectional suffix. He posits that there was no allomorphy of the suffixes in the proto-language. Roots nowadays mostly end in a vowel. This is one of the indications that a potential root final consonant may have been reanalysed as part of the inflectional ending. For Dixon, in the proto-language, roots could end in a vowel or in /y, ŋ, m, n, l, rr/ (Dixon 1980:409). He thus posits a limitation on the type of consonants occurring in coda position in verbs, much in the same way as there are restrictions in Oceanic languages. This set of consonants is also precisely the set of consonants he reconstructed that could occur in coda position in the proto-language. Dixon posits a number of sound changes that occurred at the root affix boundary, which obscured the clear segmentation of the forms, and favoured reanalysis (Dixon 1980:412-414). The final consonant of the root was understood as being part of the inflectional material of the verb forms in those forms in which it appeared, while it was deleted by sound change in some other forms.

Harvey (2008) refutes an analysis in synchrony where the thematic consonant is part of the root and shows various processes of deletion when forming a consonant cluster, on the basis that it gives inconsistent results in Guugu Yimidhirr. He then makes a distinction between open and closed classes, because for him productive systems show different patterns of change with regard to non-productive systems. In all Australian languages, there are one or two very large membership classes (often open), and a number of more restricted closed classes. He proposes, following Dixon, to reconstruct the M class and the NG class as closed classes, and all the others (L, N, Y, vowel and RR) as open classes. According to Harvey (2008) token frequency will play a major role in the evolution of closed classes, while type frequency will be more important for open classes, probably because open classes serve as models to new lexemes while closed classes never do so. In particular, Harvey states that "phonological structures that indicate internal morphological divisions tend to be eliminated" in closed class paradigms, while they tend to be maintained in open class paradigms (Harvey 2008:128-129). From the use of substantively inflected verb forms as stems for other verb

forms, Harvey concludes that the origin of inflectional class markers lies in the allomorphy of the Non-past marker, which in the proto-language presented wide variation between *-*n*, *-*ng*, and *-*r*, as shown by comparison between Gunwinyguan languages and Pama-Nyungan languages. He thus proposes a change where the TAM suffix is reanalysed as a conjugational marker. The use of non-past forms as a stem in fact varies depending on the conjugation: it is particularly used in large, open classes. The use of that form as a stem is caused by its present reference and the fact that forms with present reference always serve as citation forms in Australian languages. At the opposite, in Guugu Yimidhirr, it is the past that functions as a stem, but only in closed classes: the non-past serves as a stem in open classes. This discrepancy in the use of stems probably meant that different TAM markers were imported in other inflected forms together with the chosen stem, depending on the type of class, open or closed. The conjugation markers originate in inner affixes with TAM values, which are stem forming for other forms. Properly in fact, the process at play here is reanalysis of the conglutination type, as there is fusion of two affixes into a single, inflectional affix.

It is highly possible that the two analyses are in fact largely compatible: derivational affixes have to become bleached first to further grammaticalize, and can be interpreted as being part of the stem. Further reanalysis can ascribe part of their phonological material to the inflectional affix. The main question is whether these segments originated in meaningful elements or in base phonological material, but the process of inflectional class creation in both cases is quite the same. In favour of an occurrence through reanalysis is the fact that the process created thematic elements, which I have shown in the previous two sections on Oceanic languages, are a typical result of reanalysis. In Dixon's analysis, inflectional classes were created through secretion; in Harvey's through conglutination, which is a route that has not previously been discussed in the literature. Whatever the stance one takes on the question, inflectional classes were created in Pama-Nyungan languages through a process of reanalysis, which typically gave rise to a system of thematic elements that each characterize its own inflectional class.

6.5. Romanian nouns in -uri

All the examples of reanalysis of the secretion type adduced until now have concerned verbs. But inflectional classes can also be created by secretion in other parts of speech. In this section, I examine how secretion created a new, productive inflectional class in Romanian and Italo-Romance nouns.

Latin had a class of nouns characterized by a difference in the number of syllables in the nominative singular and in the rest of the paradigm, called imparisyllabic nouns. This was caused by the fact that the nominative singular (and accusative singular for neuter nouns) were affixless, showing a bare root. The relationship between the form of the root and the stem for other forms was obscured by sound change (vowel raising in nominative singular, and *s* > *r* in other forms). This gave forms such as *tempus* ~ *tempora* 'time'. Because the form of the root paralleled that of many masculine nouns in *-us*, there was reanalysis of that ending as an inflectional affix, giving rise to a number of new nouns with a plural in *-ora*, which previously had a plural in *-i*. This is the case with *armus~armi* 'shoulder' (M) > *armus~armora* in the fourth century (Gardani 2013:367), which shows that this class of nouns had acquired productivity, and that the ending *-ora* had been reanalysed as a plural marker on its own, from parts of the original root and part of the inflectional ending.

When the gender system of late Latin changed and ousted the neuter forms, most neuter nouns were reassigned unambiguously to either the masculine or the feminine gender. Imparisyllabic nouns, though, were sometimes assigned to masculine in the singular and to feminine in the plural, because their endings exhibited the form characteristic of the main productive class for these genders: *-a* for feminine singular of class 1, *-us* for masculine singular of class 2. Maiden (2016c) argues that this is an element in favour of an analysis whereby *-or-* is considered as part of the lexical root, because *-a* is segmented as realizing feminine gender. This gender alternation is characteristic of the large class of nouns in Romanian that show the Ø~*-uri* alternation. There has been a subsequent extension of this pattern to a large number of nouns which did not present the

alternation historically in Romanian. This means that the initial reanalysis of *-uri* as a plural form has been extended to novel lexemes as a productive pattern. This is evidence that an element which was initially part of the root (*-or-* or *-ur-* in Romanian) has been reanalysed as part of the inflectional ending, or at least as a thematic element appearing in the plural only. If, following arguments in Maiden (2016c) one recognizes that in the history of Latin and its Romance continuants, the element *-or-* or its continuants have been sometimes ascribed to the root, sometimes to the inflectional affix, and sometimes to neither, one can analyse it as a thematic element that only signals the inflectional class those nouns belong to. This is again an expected result of the creation of an inflectional class through reanalysis.

6.6. Conclusion

Reanalysis of the secretion type has been recognized in the literature as a source of inflectional classes. This chapter has provided more examples of such a phenomenon, proving that it is not a minor source for the creation of new inflectional classes. It is attested in a number of Austronesian languages, as well as Pama-Nyungan languages of Australia (on Dixon's analysis), and the Romance languages. Most examples adduced concern verbs, but the Romanian example has shown that it is not limited to that part of speech, and that nouns can also undergo secretion which gives rise to new inflectional class distinctions. Finally, this chapter shows that there is a peculiarity in classes created through reanalysis. All the examples I have analysed in this chapter exhibit classes characterized by the presence of a thematic element. As such, reanalysis in those cases does not exactly proceed as expected: a segment of the root is not reanalysed as part of the affix strictly speaking, but is analysed as a new formative which only flags the inflectional class the lexeme belongs to. Such thematic elements were not present in the inflectional classes resulting from other sources. One has thus to recognize reanalysis as the major source of inflectional classes with thematic segments. Conversely, it means that if an inflectional class system exhibits thematic

elements, its origins probably lie in reanalysis.

7. From alienability distinctions to inflectional classes

This chapter examines the evidence that a source of inflectional classes is alienability distinctions in possessive paradigms, and that this is particularly the case in South Halmahera-West New Guinea languages (section 7.4) and non Austronesian languages from Papua (section 7.5), although the phenomenon also occurs in other language families around the world (section 7.6). I show that the rise of inflectional classes is highly correlated with head marking in nominal possessive paradigms. This is a specific case where a system of inflectional classes originates in another system presenting an already existing system of classes based on semantics.

7.1. Alienability distinctions in possessive paradigms

Alienability distinctions are a form of nominal classification arising in possessive paradigms (Nichols & Bickel 2013). They typically oppose two constructions for the marking of possession, one said to be 'inalienable', the other 'alienable'. Those oppositions are lexical classifications concerning the head noun of the construction, that is to say the possessum, that which is possessed. An opposition between just two classes is by no means the only possible split, and many languages present more classes (see below). Similarly, the formal means for expressing those classes are not always uniform, and many language descriptions set up an opposition between two classes by subsuming other possible subclasses. Alienability distinctions are found on all inhabited continents, but they are particularly prevalent in Austronesian languages, non-Austronesian languages of Papua, Australian languages, and North American languages.

Inalienably possessed nouns typically include body parts or kin terms, though the type of nouns belonging to the inalienable class varies between languages (Nichols 1988:572). Alienable nouns are simply the remainder of nouns in the language, for those languages which distinguish two classes. Although the inalienable class typically includes kin terms and body parts, it is never a

unitary class defined by strict rules of membership: most often, not all kin terms and body parts will belong to it, or it will have further members chosen particularly among those words that express a part-whole relationship (Nichols 1988:572). For example, in Eastern Pomo, the inalienable class comprises a closed, unproductive subset of kin terms (McLendon 1975:92, 108). Inalienables are marked in Eastern Pomo affixally, as in Example 1, while alienables are marked through a genitive marker on the possessor, as in Example 2:

1) wí-bayle
   1SG-husband
   "my husband"
2) wáx šaˑri
   my.GEN basket
   "my basket"

Alienability distinctions cannot generally be described in terms of precise semantic distinctions (see Nichols 1988:568-576), even for those systems which have been said to be semantics based. The labels tend to indicate that there is a distinction between more permanent, or inherent types of possession, and less permanent types. In a way, the opposition seems to be between inherent possessive or part-whole relationship, and cases of ownership. Thus inalienable possession expresses the more permanent, inherent, type of possession: a head or arm is inherently possessed, without that possession being amenable to changes in ownership. Similarly, kinship relations are perceived as being permanent and non-modifiable: one's father remains one's father. Conversely, alienable possession is claimed to represent the other extreme, things that are possessed or owned, but whose ownership can change: my axe can become someone else's if I sell it or if I give it to someone. This has sometimes also been paralleled with the distinction between relational and non-relational nouns, where relational nouns are those whose lexical semantics include a relationship between two entities, and not simply the denotation of an entity: a sister designates the relationship between two entities (one is always a sister of someone), not simply a single entity. The problem

with this view is that the semantics of each class is by no means identical crosslinguistically, nor is it easily definable in terms of semantic domains for any given language. The major problem is the impossibility of defining common semantics for inalienables in a given language, for what should be a morphosemantic feature. If all kin terms and body parts were marked for inalienable possession in a language, the class would be definable from semantics, not only from formal marking of possession. But in most languages this is not the case. Thus for example, in Eastern Pomo only some kin terms are marked for inalienable possession. This does not mean that in that language only some of the kin terms are perceived as having a permanent relationship of possession, whilst others are not. Such explanation would merely be post hoc. For example there is generally no single denominator to be found between terms that are considered alienable or inalienable: such generalisations could be affinal versus blood relationships, or nuclear family versus extended family. This simply means that some lexical items, that happen to come from similar lexical groups, mark possession in a peculiar way (see Nichols 1988 for more developments on this point). In Eastern Pomo, such nouns mark possession with head marking, affixally. But even there, one should note that there are two main patterns of inalienable marking (McLendon 1975:114; see below).

In any case, there is a clear asymmetry between the inalienable and the alienable class. In particular, most languages oppose a closed class of nouns belonging to the inalienable type of marking of possession, and an open, productive class or classes exhibiting the alienable type of marking of possession. Furthermore, the number of nouns in the inalienable class is generally quite small.

The distinction between alienable and inalienable nouns can take many forms in possessive paradigms. Typically, inalienables will be head-marking, and marked affixally, most often through the affixation of a person and number agreement marker on the head noun, which agrees with the possessor. An example comes from Warndarang (Australia, Heath 1980:28-29, cited in Nichols & Bickel 2013), where some kin terms and other nouns take head-marking in the form of a possessive

prefix (Example 3), while the remainder of nouns in the language take an independent genitive pronoun to express possession (Example 4):

3) ng-baba
   1-father
   'my father' / 'our father'

4) wu-radburru ngini
   NCM-country 1SG.GEN
   'my country'

Nichols (1988:578) concludes from a large sample of North American languages that there is a strong correlation between head-marking of possession and alienability distinctions. In her sample, only one language, together with some Polynesian languages, makes a distinction in alienability although being strictly dependent marking. Although marked through affixation, inalienables are not a homogeneous class in all languages, either morphologically or semantically (see section 7.4. to 7.6 below). Alienables on the other hand tend to be marked either through affixation, as in Magey Matbat (Austronesian, South Halmahera-West New Guinea, Example 5), or by an indirect, periphrastic construction involving the affixation of agreement markers with the possessor to a support word, a possessive marker, which is sometimes itself a possessive classifier, as in Fijian, or in Eastern Pomo (Example 2 above).

5) Magey Matbat (van den Berg 2009:336)
   a. fa-m
      husband-2SG
      ‘your husband’
   b. aw-wa^{3}ŋ
      2SG-canoe
      ‘your canoe’

Alienables can also be marked through a specific possessive word or genitive marker, as in

Warndarang (Example 4 above), or in Eastern Pomo, most typically dependent marking.

Those forms have sometimes led researchers to distinguish between direct and indirect possession, which is a way to indicate both the degree of permanence of the possessive relationship, and the formal marking used to express it. Thus direct possession will be marked through head marking, directly on the head possessed noun, and affixally. Indirect possession wil make use of one or more possessive markers, which can typically take the form of a possessive classifier. This is in particular the case in Oceanic languages (Lichtenberk 1985, 2009). Lichtenberk exemplifies the typical Oceanic type of constructions with Manam, and says that it has two main types of constructions, one corresponding to inalienables, in direct head marking, as in Example 6 (Lichtenberk 2009:149), the other type of construction attaching agreement markers to a possessive classifier, as in Examples 7 and 8 (Lichtenberk 2009:249-250) and corresponding to alienables:

6) ara-gu
   name-1SG:POSS
   'my name'

7) pera ʔana-gu
   house POSS.CLF-1SG:POSS
   'my house'

8) asi ne-gu
   bushknife POSS.CLF-1SG:POSS
   'my bushknife'

Finally, some languages distinguish obligatory and general possession, a distinction sometimes equated with inalienability. In some languages, some terms, often similar to the items found typically in the inalienable class, cannot appear without possessive marking. The notion is not identical to inalienability, but it plays a role in the first appearance of such distinctions in a language, as will be shown in the next section.

Alienability distinctions are comparable to gender systems with at least partially a semantic

assignment of nouns to genders. A large number of items do not obey such a semantic assignment, but there is in language after language a core of semantically similar lexical items belonging to the inalienable class. It is certainly possible that some of those systems may be truly semantically based, when they only oppose two classes, and when these classes are perfectly defined by lexical semantics: for example for those languages where all kin terms and body parts belong to the inalienable class. Similarly, those languages where recategorisation is possible depending on the contextual meaning of the word are more truly semantically based: for example when a severed head is reassigned to the alienable class. This chapter is mostly concerned with systems where such recategorisation is impossible, and where the assignment of nouns to a given class seems more lexical than purely semantic.

7.2. Commonalities between alienability distinctions and inflectional classes

There are two major commonalities between alienability splits and inflectional classes: the fact that they concern inflection and the fact that they are both lexical classes.

First, only those alienability splits that involve inflectional morphology can be profitably compared with inflectional classes. In effect, this will correspond to all languages for which at least inalienables are marked through inflection, which corresponds to languages that have head marking of possession through person and number affixes affixed to the possessum head. For those languages, possession can be analysed as an inflectional category, even though in some cases the marking can take the form of a periphrasis (for alienables). If possession is an inflectional category, distinguishing possessor agreement in addition to the meaning of possession, then formal differences in the expression of the category for different lexemes have to be considered inflectional classes. Thus in Eastern Pomo, two patterns of inflection can be recognized for the so-called inalienable class, remaining nouns belonging to a periphrastic, alienable class. Similarly, in Magey

Matbat (Example 5 above), both classes are marked inflectionally, and correspond to inflectional classes which are weakly linked to semantic classes for the inalienable class. It is thus a case where the rise of inflectional classes can be correlated with head marking.

The second major point in common is the fact that both inflectional classes and alienability classes are lexical classes (Nichols and Bickel 2013). They generalize over discrete sets of lexemes. Each lexeme has to be assigned to one of the available classes in the grammar of the language. As Nichols & Bickel (2013) put it, "possessive classification is not a semantic or grammatical category but a purely lexical classification of nouns. Meaning is involved in that there is usually a default or open class and a specified or determinate class with a semantic common denominator shared by most but not all members of the class and also sometimes found in non-members". In most cases, no lexical item can be recategorized to another class. When this happens, it generally involves a very special kind of construction, for example creating a compound which is not inflected according to the same possessive class. Such cases can be simply analysed as cases of word formation, whose result is a word belonging to a different inflectional class.

Just like inflectional class systems, alienability split systems tend to oppose one or more closed class of irregulars, with very few members, and an open, productive class of inflection. This division parallels the division between the closed class of inalienables, with few members, and the open, productive class of alienables.

One can thus make a tentative generalization: when a language distinguishes more than one type for the marking of possession (when a language has an alienability split), and at least one construction realizes possession inflectionally on the head noun (head marking), that distinction is one of inflectional class.

Most alienability distinctions should thus be considered as the way possession is realized differently for different lexical items. I show in Sections 7.4., 7.5. and 7.6. that it is indeed the case for a number of languages from various families, including Austronesian, Papuan and other languages. The interesting point in considering such alienability splits as inflectional classes is that

it makes a number of predictions that seem to hold true from what is stated in the literature.

The first prediction is that recategorization of nouns from one class to the other should be impossible in synchrony. Recategorization is expected to take place in diachrony, which amounts to some lexical items changing inflectional class. A number of languages allow such recategorization, but through specific devices that have to do with the derivation of a new lexeme belonging to a different inflectional class for possession. The less recategorization is possible, the closer an alienability split will be to an inflectional class system.[46]

The second prediction is that classes do not have to be semantically well defined. In section 7.1. I have shown that in most cases the inalienable class is not perfectly definable through the lexical semantics of its members. Thus, it is not possible to invoke a feature "kin term" or "body part" to define the whole class, because not all such terms belong to the inalienable class, or because some other lexemes not definable in such terms do belong to it. Such a prediction holds true for systems of alienability split involving inflection.

The third prediction is that in such systems one can have as many classes as imaginable. In theory, the maximal number of inflectional classes could be equal to the number of lexemes in a given word class. If the systems were not of inflectional classes but of semantically based classes, the number of classes would be limited by semantic considerations. In a system of inflectional classes, the number of classes is not limited a priori. French verbs for example have been described as having a very high number of inflectional classes, some counts going up to 72 (Stump 2016:87). Furthermore, one could expect the different classes to be organized like existing systems of inflectional classes, for example they might show heteroclisis, or be organized as inheritance hierarchies with default markings for some of the cells. This is indeed what is observed. Some languages show heteroclisis in possessive paradigms, for example Biak and Ambai (see below 7.4.).

46 Another type of alienability distinctions involves semantically systematic and opposed meanings under possession such as inherent possession "my skin" and alienable possession "my (animal) skin", the skin I use to make clothing, for example (see Nichols 1988:565-566). This type of opposition could be considered as simple conversion of lexemes, a kind of derivation that changes the inflectional class of the noun without affecting its phonological shape. In a number of languages, there is special morphology attached to this change. In Biak for example, it will often be through compounding.

Others show default marking for some cells, such as Irarutu (see below 7.4.3.)[47].

Finally, in an inflectional class system, one very often sees at least one productive, regular class, and to have other classes that are more limited in terms of membership. Those classes, if showing more irregularity, should comprise very frequent lexical items. This is what is seen in alienability split systems. Although canonically the productivity of inflectional classes should not be correlated with meaning, there is a trace in inflectional class systems of possessive paradigms of some semantic grounding for the less productive classes at least; this is because such frequent items tend to come from similar classes such as body parts and kin terms.

7.3. Acquiring alienability distinctions

A number of explanations have been given for the origins of alienability distinctions. It is important to understand how alienability distinctions arise in language, because it can explain how some inflectional class distinctions can arise from them. In some cases, scholars have proposed that alienability distinctions can be acquired through prolonged contact inside a linguistic area. This is particularly the case for Austronesian languages spoken on and around the island of New Guinea (Donohue & Schapper 2008). Others have appealed to the notion of iconicity to explain the rise of alienability distinctions, whereby more permanent and direct types of possession are marked affixally, whereas less permanent types of possession and cases of ownership are marked indirectly (Haiman 1985). In opposition to iconicity explanations, the most widely accepted explanation relies on questions of high frequency of possessed items and regrammaticalization of possessive marking to explain the distinction between two or more types of marking (Nichols 1988; Dahl & Koptjevskaja-Tamm 1998; Haspelmath 2008, 2017). I will rely on this last type of explanation here. Iconicity explanations are problematic because not all languages mark the arguably more permanent possession affixally, and some languages still distinguish two classes while marking nonpermanent

[47] Another element, but less diagnostic, is the prevalence of suppletion in such paradigms (Baerman 2014).

possession affixally as well. An iconicity explanation in fact only describes some of the situations found, while a regrammaticalization hypothesis can account for all attested systems, and is also attested in systems where we have historical documentation, such as the distinction made in Catalan for example.

Dahl & Koptjevskaja-Tamm analyse the rise of alienability splits in language. For them, alienability splits appear through a process of grammaticalization of a new, alienable construction. The construction must somehow be blocked for inalienables, mostly kin terms and body parts. They show that frequency and obligatorification of the possessive construction for such items plays a role. In particular, they show that a number of languages have articleless kin terms opposed to other types of lexemes appearing with an article, which is a type of alienability split common throughout Europe. Because all languages that have a split possessive paradigm with lexical possessors also have it with pronominal possessors, it is possible to conclude that the split first originates in pronominal possession, and they conclude that "an expanding possessive construction must encroach on the territory of pronominal possession for an alienability split to arise" (Dahl & Koptjevskaja-Tamm 1998:48). For them this is the only way in which alienability distinctions are created.

Haspelmath (2008, 2017) proposes an explanation of alienability splits in terms of frequency of use, and analyses two main paths for their creation, the rise of an innovative pattern for alienables which is blocked for more frequently possessed items, and also adds another path, the differential phonological reduction of markers for more frequently possessed items, for which he gives evidence from old Italian and Lancashire English. The first development has to do with what he calls differential phonological reduction. It has to do with the common shortening of very frequent expressions in language, particularly through differential sound changes that lead to the greater reduction of frequent forms. In such cases, "a shorter inalienable pattern arose by phonological reduction from a fuller pattern, while the corresponding alienable pattern does not show the same reduction" (Haspelmath 2017). Examples given by Haspelmath (2017) include the following

examples from old Tuscan (9) and Lancashire English (10):

9) a. moglia-ma < Lat. mulier mea 'my wife' (inalienable)

b. terra mia < Lat. terra mea 'my land' (alienable)

10) a. m[ɪ] brother (inalienable)

b. m[aɪ] football shoes (alienable)

Haspelmath does not point out that there is also a syntactic difference in his Tuscan examples, in that Example 9b must take a definite article while this is impossible for nouns marked with an enclitic. Italo-Romance inalienables also include nouns which are not sematically unitary and part of the kinship term system, such as *casa-ma* 'my house'.

Note that this path tends to create new head marking for typically inalienable nouns. One can easily find other examples of the same path of change in the Romance languages, where it is attested at least in Romanian, and in central and southern Italian dialects. Romanian for example uses enclitic pronominal forms on kin terms and some close relation nouns (Maiden 2016:117). These show a reduced form of the possessive adjective, but some items also show a reduced form of the noun. Examples such as 11 and 12 (Maiden 2016:117) show a clear case of differential phonological reduction for some high frequency kin terms, where *mama* is reduced to *mă*, and *fratele* to *frac*, when marked with the enclitic possessive:

11) mă-sa vs mama sa

mother-his mother his

'his mother'

12) frac-su vs fratele său

brother-his brother his

'his brother'

The second path has to do with the "differential inhibition of an expanding expression". It is essentially the path described by Dahl & Koptjevskaja-Tamm (1998), whereby a new, periphrastic construction starts being used for those nouns where a reinforcement of the possessive relationship

is felt necessary, which tends to be for typically alienable contexts, in cases of ownership. A split will arise when the development of the new construction is blocked for lexical items which are most often possessed, that is typically inalienable lexemes such as kin terms and body parts, because of the high frequency of earlier forms. Examples include Egyptian and Coptic, as well as the Maltese example given by Dahl and Koptjevskaja-Tamm (1998), as well as the Biak alienable construction which will be analysed in Section 7.5.

The two paths exemplified by Haspelmath are by no means mutually exclusive, and I show in Section 7.5. that both were used in the rise of inflectional class marking in possessive paradigms in the Austronesian language Biak. The two possible paths analysed by Haspelmath and Dahl & Koptjevskaja-Tamm correspond to the two possibilities evoked in a seminal paper by Johanna Nichols (1988). They also combine two of the main possibilities for the creation of inflectional classes, mainly sound change and grammaticalization, with the added specificity in the case of alienability distinctions that it is linked to the semantics of some terms, and the frequency of their appearance in a possessive construction.

7.4. Possessive paradigms in South Halmahera-West New Guinea show inflectional classes

7.4.1 Evidence from Biak

Biak has been described as showing a distinction between alienable and inalienable possession (van den Heuvel 2006; Mofu 2005, 2008; Steinhauer 2003, 2005). Biak shows a range of head marking constructions where possession is marked affixally for items that typically tend to be included in an inalienable class, mainly kin terms and body parts. As in most other languages, not all kin terms and body parts take this type of marking. Other lexemes are marked periphrastically through a marker originally deriving from a relative clause involving a possessive verb that agrees with both the possessor and the possessee (Bach 2017). The authors who distinguish for this

language alienable from inalienable possession in fact oppose those items that are head-marked (so called inalienables), and those that are not (so called alienables). In doing that, they in fact conflate a large number of patterns into an artificial inalienable class: Biak in fact presents different classes with a semantic link, distinguishing two classes for body parts, one for kin terms, and one for locational nouns, in addition to a number of heteroclite patterns for some kin nouns.

7.4.1.1. Alienable pattern

The so called alienable construction in Biak is a relatively recent innovation in the language dating from proto-Biakic, since it is not attested in other South Halmahera-West New Guinea languages (Bach 2017). It is a postnominal marker agreeing in gender, number, and givenness with the possessee and in person, gender and number with the possessor:

13) Yohanes rum byedia
Yohanes rum b<y>e-dia
John house <3SG>POSS-DET.SG.DEF
'John's house'

14) ben mu-v-an-sya si-ba
pig 2SG-POSS-GIV-DET.PL 3PL-big
'Your (DU) pigs [that we already mentioned] are big'

The alienable marking is periphrastic: possession is marked through a separate word. This word originated in the grammaticalization of an object relative clause, roughly from "house, I have it" to "my house" (Bach 2017), which explains why the marker is made up of a verbal root *ve* with possessor prefixal agreement and a closing determiner agreeing with the possessee, as in Table 89 (van den Heuvel 2006:230):

	SG	DU	PCL	PL
1INCL	(a)ye=dya	ku-ve=dya	ko-ve=dya	
1EXCL		nu-ve=dya	(i)nko-ve=dya	
2	be=dya	mu-ve=dya	mko-ve=dya	
3ANIM	v<y>e=dya	su-ve=dya	sko-ve=dya	s-e=dya
3INAN				n-be=dya

Table 89. Alienable possessive pronoun in Biak (van den Heuvel 2006:230)

7.4.1.2. Inalienable patterns: body parts

The inalienable class in Biak in fact covers a number of different types of marking for different lexical items. According to van den Heuvel (2006:232), "Inalienable morphology, then, is found with a number of kinship terms, with a number of bodypart terms and with a number of locational nouns or nouns referring to parts of wholes. Each of these groups has its own paradigm." What is interesting here is the fact that there is not a single, unitary paradigm for inalienables, but a series of paradigms partially linked to the semantics of the lexical item. The number of items entering those so called inalienable constructions amounts to about thirty lexemes: eleven body part terms, ten kin terms, and about a dozen locational nouns and other items (van den Heuvel 2006:232-251).

Body part terms inflect according to two main inflectional classes.[48] The first pattern concerns, mainly, body parts which normally occur in pairs, the second body part terms that take inalienable marking (van den Heuvel 2006:238-242; Mofu 2008:57). The two classes differ only in the singular. Both are formally marked only by suffixes in the singular, and by both prefixes and suffixes in non singular numbers. In both classes, 1SG and 3SG cells are syncretic. The paired body parts pattern does not distinguish between singular and dual for the possessed, such that *we-si* means both 'my leg' or 'my legs'. The paradigm for this inflectional class is given in Table 90 (van den Heuvel 2006:240; Mofu 2008:57):

48 Most body parts also have a compound counterpart which inflects according to the alienable pattern. One must still consider those as different lexemes, potentially with a slight difference in meaning. Thus a number of body parts can appear in compound with the word *kor* 'bone', such as *vukor* 'head' instead of *vru* 'head'.

POSSESSOR	FORM
1SG	we-si
2SG	we-msi
3SG	we-si
1DU.INCL	ku-we-sna
1DU.EXCL	nu-we-sna
2DU	mu-we-msna
3DU	su-we-sna
3PCL	sko-we-sna
1PL.INCL	ko-we-sna
1PL.EXCL	nko-we-sna
2PL	mko-we-msna
3PL	si-we-sna

Table 90. Paradigm for we *'leg' in Biak*

The nouns following this inflectional class include *mka* 'eyes', *kna* 'ears', *vra* 'hand, arm' and *snoni* 'nose' (Mofu 2008:57).

The second pattern for body part terms only differs in the singular, as shown in Table 91:

POSSESSOR	*we* 'leg'	*vru* 'head'
1SG	we-si	vru-ri
2SG	we-msi	vru-mri
3SG	we-si	vru-ri

Table 91. Different paradigms for body parts in Biak

Nouns such as *sne* 'stomach', *sva* 'mouth', *kro* 'arse', *sasu* 'neck' follow this second inflectional class. This is also the case for terms not referring to body parts such as *sno* 'name' or the locational nouns *fadu* 'middle' and *do/don* 'inside' (van den Heuvel 2006:238). Other body part terms, such as *andar* 'forehead', *ramar* 'tongue' or *sus* 'breast' follow the alienable pattern. The classes, although they show some commonality in terms of lexical semantics of their members, are thus not fully definable by such semantics.

7.4.1.3. Inalienable patterns: kin terms

The pattern of inflection for kin terms involves suppletion for 1SG. The remaining inflection is regular for all terms following this inflectional class. Steinhauer (2003:16-17) gives full paradigms for both *awin/sna* 'mother' and *apus/kpu* 'grand-child', at least for singular and dual possessed. He seems to indicate that paucal and plural possessor are also possible (Table 92). This may reflect a

difference in dialect, as Mofu (2008:57) indicates that *kpu* is a heteroclite, and van den Heuvel (2006:243) gives a very different, heteroclite, paradigm.

	POSSESSED	
POSSESSOR	SG	DU
1SG	awin-i	awin-su
2SG	sna-m-i	sna-m-su
3SG	sna-ri	sna-su
1DU.INCL	ku-sna-ri	ku-sna-su
1DU.EXCL	nu-sna-ri	nu-sna-su
2DU	mu-sna-m-i	mu-sna-m-su
3DU	su-sna-ri	su-sna-su
3PCL	sko-sna-ri	sko-sna-su
1PL.INCL	ko-sna-ri	ko-sna-su
1PL.EXCL	nko-sna-ri	nko-sna-su
2PL	mko-sna-m-i	mko-sna-m-su
3PL	si-sna-ri	si-sna-su

Table 92. Inalienable kin term pattern in Biak: awin */* sna *'mother' (Steinhauer 2003:16-17)*

Lexemes following this pattern of inflection include *kamam/kma* 'father', *awin/sna* 'mother', *imem/me* 'cross uncle', *kafnom/fno* 'cross sibling's child', *apus/kpu* 'grandparent, grandchild' (van den Heuvel 2006:244).

Some kin terms also inflect according to the alienable pattern. Among those, one can single out *romawa* 'son', *inai* 'daughter', and *naek* 'parallel sibling (same sex sibling)', which are not peripheral members of the system of kinship in Biak, but simply follow a different inflectional class (van den Heuvel 2006:246).

7.4.1.4. Inalienable patterns: locational nouns and other nouns

Locational nouns are also grouped with inalienables by descriptive grammars. They show a very interesting pattern of inflection in Biak, involving defectiveness for their possessive paradigm. Thus only the forms in Table 93 exist for the lexeme *bo* 'upside' (van den Heuvel 2006:250), while grey cells indicate impossible realizations (this gap having no specific semantic motivation):

	SG	DU	PCL	PL.ANIM	PL.INAN
1					
2	bo-mri				
3	bo-ri	bo-nsu	bo-nsko	bo-nsi	bo-nna

Table 93. Locational noun in Biak: bo *'upside' (van den Heuvel 2006:250)*

Nouns following this defective pattern include *bo* 'upside', *do* 'inside', *fadu* 'middle', *bav* 'downside', *rawn* 'front part of a canoe', *fúwar* 'bottom part of a canoe', *wurn* 'back part of a canoe', and *andír* 'side'.

In the end, the so called inalienable class is in fact made up of four different inflectional classes that are partially correlated with the lexical semantics of its members. In particular, the assignment of a lexical item to an inflectional class depends on the semantics of the term, whether it refers to a body part, a kin term or is a locational noun, but such information is not sufficient to predict the precise inflectional class to be assigned. In addition to that, some body part and kin terms also inflect according to the periphrastic alienable pattern.

7.4.1.5. Heteroclisis

In addition to these patterns, a certain number of items show heteroclisis, a situation which is expected only if the marking of possession involves inflectional classes, and not a semantically based split. In effect, such items inflect for part of their paradigm following the so called alienable pattern, and for the other part following one of the inalienable patterns. This amounts to saying that such items are heteroclites where parts of the inflectional paradigm are expressed periphrastically. Such items include *swa* 'spouse' and *srar* 'cross-sibling', *na* 'tooth' and *vrampin* 'fingers' (this last item a compound between *vra* 'arm' and a cranberry morph).

According to van den Heuvel (2006:243-245) all kin terms show heteroclisis between a so-called alienable and inalienable pattern, in addition to some kin terms showing suppletion for 1SG. This may reflect a change in progress, as two of the nouns can also be inflected fully according to the alienable pattern. The main pattern of heteroclite inflection is as shown in Table 94 for *imem/me* 'cross uncle' (van den Heuvel 2006:243). The grey zone in the table represents the section of the paradigm that follows an alienable, periphrastic inflectional class:

POSSESSOR	SG	DU	PCL	PL
1SG	imem-i	imem-su	imem-sko	
2SG	me-mi	me-msu	me-msko	
3SG	me-ri	me-rsu	me-rsko	
1DU.INCL				
1DU.EXCL				
2DU				
3DU				
3PCL				
1PL.INCL				
1PL.EXCL				
2PL				
3PL				

Table 94. Heteroclite inflection for imem */* me *'cross-uncle' in Biak (van den Heuvel 2006:243)*

However, according to Mofu (2008:61), only the forms for singular possessee with a dual, paucal or plural possessor are realized according to the alienable pattern. This probably reflects dialectal differences in the variety of Biak chosen by Mofu (multilectal) and van den Heuvel (Biak Wardo).

The second pattern of heteroclisis comes quite close to the first one, except that there is no suppletion for 1SG, which instead is realized according to the alienable inflectional class. The same lexical items follow different inflections in Mofu's and van den Heuvel's descriptions. For Mofu (2008:61), this pattern only involves alienable marking for 1SG possessor, while for van den Heuvel (2006:245), only 2SG and 3SG are inflected according to an inalienable pattern. This pattern includes *srar* 'cross sibling', *swa* 'spouse' (van den Heuvel 2006:245) as well as *mambanyo* 'father in law' and *imbanyo* 'mother in law' (Mofu 2008:57). It seems that some dialects also place in that inflectional class *kpu* 'grandchild' (Mofu 2008:57). In any case, what matters here is that these lexemes show heteroclisis, or at least that their inflectional class originated in heteroclite paradigms, notwithstanding their dialect: part of their inflectional paradigm is inflected following one inflectional class (the alienable class), while the remainder of the paradigm is inflected according to another inflectional class (one of the inalienable classes).

The last pattern of heteroclisis opposes 1/2SG inflected according to the paired body parts inflectional class, and the rest of the paradigm inflected according to the alienable class. Nouns

inflecting according to this pattern include *na* 'teeth', *vrampin* 'fingers', and *napirm* 'cross cousin' (van den Heuvel 2006:245; Mofu 2008:59).

7.4.1.6. Conclusion on Biak possessive inflectional classes

To sum up, more than half the lexical items belonging to the so called inalienable class in fact show heteroclisis between one inalienable pattern and the periphrastic, alienable pattern. The form of heteroclisis takes two different paradigms. This probably means that those items form in synchrony a separate inflectional class in Biak, but that historically that class arose from heteroclisis. The other half of so-called inalienable items shows four different patterns of inflection. This shows that the inalienable class is in fact made up of a number of closed membership, older, inflectional classes, and similarly, with the number of heteroclite lexemes, that the alienable pattern is yet another, open membership and productive, inflectional class. This point is further proven by the fact that heteroclisis is a common occurence in systems of inflectional classes (Stump 2006; Stump 2015), be it in synchrony or in diachrony (Kaye 2015).

It is possible to infer how inflectional classes arose for possessive paradigms in Biak. The alienable class is a more recent development than the inalienable classes, having grammaticalized from what was originally a relative clause involving a verb of possession (Bach 2017). The various inalienable classes seem to have grammaticalized earlier, and involved classifiers. At least two possessive classifiers have been reconstructed for proto South Halmahera-West New Guinea, **ri* as a general classifier, and **na* as an edible possession classifier. It is possible that there were more. The various inalienable classes show the fusion into one word of both bound pronominal forms agreeing with the possessor, and of reflexes of classifiers **ri*, **na*, and pronoun **si*. Thus the pattern mainly used for paired body parts includes a reflex of classifier **na*, for all non singular numbers:

15) si-we-s-na
3PL-leg-3PL-POSS.CL
'their legs'

Kin terms mostly include reflexes of pronouns having fused first as clitics, then as affixes, with

the head noun of the construction. Biak shows both types of acquiring new possessive distinctions: inalienables show a progressive phonological reduction of a structure first involving a pronoun and a noun to yield bound forms; alienables show the rise of a new construction that was blocked for items most commonly possessed (i.e. inalienables). Finally, the heteroclite patterns arose after the new pattern was in place, and show a progressive entrenchment of that construction in inalienables.

7.4.2 Evidence from Ambai

Similar evidence to that from Biak can be found in the South Halmahera-West New Guinea language of the West Yapen subgroup Ambai (Silzer 1983; Karubaba 2009). Ambai is also described as having an alienability distinction (Silzer 1983:123), but it only has one main inalienable pattern. Only some kin terms and some body part terms are inflected according to the inalienable pattern. It can nevertheless be shown that these distinctions in Ambai are in fact inflectional classes.

7.4.2.1. Alienable marking in Ambai

Alienable marking in Ambai is marked through a preposed possessive marker *ne* which is marked for possessor agreement (Silzer 1983:124), through suffixes in the singular, and prefixes in non-singular numbers. The possessive marker has the paradigm presented in Table 95 (Karubaba 2009:26):

	SG	DU	TR	PL
1.INCL	ne-hu	tu-ne	to-ne	ta-ne
1.EXCL		a:u-ne	anto-ne	ame-ne
2	ne-mu	mu-ne	munto-ne	me-ne
3	ne	u-ne	co-ne	e-ne

Table 95. 'Alienable' possessive markers in Ambai (Karubaba 2009:26)

Most lexemes in the language follow this pattern of marking possession, including a number of body part and kin terms (Silzer 1983:124).

7.4.2.2. Inalienable marking in Ambai

Inalienable marking in Ambai concerns a closed set of kin terms and body part terms, but not all

such nouns are inflected according to this pattern. Nouns are marked for possessor agreement through affixes. In non-singular numbers, in addition to a specific person and number prefix, a non-singular suffix *-mi* is added. The pattern of inflection follows the system presented in Table 96 (Silzer 1983:88), where X represents the base noun. The exponents used for this marking are related to those used on the alienable possessive marker (Table 95) to a large extent.

	SG	DU	TR	PL
1.INCL	X-ku	tu-X-mi	to-X-mi	ta-X-mi
1.EXCL		au-X-mi	anto-X-mi	ame-X-mi
2	X-mu	mu-X-mi	munto-X-mi	me-X-mi
3	X-n / -na	u-X-mi	ito-X-mi	e-X-mi

Table 96. Inalienable marking in Ambai (Silzer 1983:88)

There is a minor distinction for 3SG between two subclasses: kin terms take *-na*, while body parts take *-n*. Karubaba (2009:25) describes a similar system, except that 1SG is *-hu*. Nouns following this inflectional class include *awe* 'foot', *nu* 'head', *wara* 'hand', *tara* 'ear', *tafere* 'tongue', and *ene* 'abdomen' for body parts; and *tama* 'father', *roro* 'cross-sibling', *ina* 'mother', *tafu* 'grandparent', and *nio* 'parent in law' for kin terms (Silzer 1983:89).

7.4.2.3. Heteroclisis

Like Biak, Ambai shows a number of heteroclite lexemes. Those all follow the same pattern, whereby singular forms inflect according to the inalienable class, and non-singular forms inflect according to the alienable class. All the members of that heteroclite pattern refer to body parts. According to Karubaba (2009:26-27), some 28 terms inflect according to this pattern, which is exemplified for *ure* 'eyes' in table 97 (Karubaba 2009:26):

	Possessive pronoun ('alienable')	*ure ‘eyes’
1SG	ne-hu	ure-hu
2SG	ne-mu	ure-mu
3SG	ne	ureng
1DU.EXC	a:u-ne	a:u-ne ureng
1DU.INC	tu-ne	tu-ne ureng
2DU	mu-ne	mu-ne ureng
3DU	u-ne	u-ne ureng
1TRIAL.EXC	anto-ne	anto-ne ureng
1TRIAL.INC	to-ne	to-ne ureng
2TRIAL	munto-ne	munto-ne ureng
3TRIAL	co-ne	co-ne ureng
1PL.EXC	ame-ne	ame-ne ureng
1PL.INC	ta-ne	ta-ne ureng
2PL	me-ne	me-ne ureng
3PL	e-ne	e-ne ureng

Table 97. Heteroclite inflection in Ambai: ure *'eyes' (Karubaba 2009:26)*

The fact that there are heteroclite lexemes mixing the two main patterns of inflection for possession proves that this is an inflectional class system, and not a semantically based system. This is the case even if heteroclisis is only a diachronic phenomenon having given rise to a new, separate inflectional class: in such a case original heteroclisis is strong evidence for the morphologization of the class distinction in the language. The interesting point in Ambai is that it shows that even a language that only has one alienable pattern and one inalienable pattern in fact exhibits an inflectional class system. This means that languages that do not show heteroclisis, but show two classes of inflection for possessive paradigms where the classes are not fully defined by precise semantics (e.g. all kin terms are of class one), can in fact be considered as having a system of inflectional classes where the assignment of lexical items to a class is partially correlated with the lexical semantics of that noun.

7.4.3 Evidence from Irarutu

Irarutu is a South Halmahera-West New Guinea language spoken on the Bomberai peninsula (van den Berg & Matsumura 2008; Jackson 2014). It has been described as presenting an opposition between alienable and inalienable possession, although descriptions do recognise that there are in

fact many more classes morphologically. In this section, I show that Irarutu in fact presents a system of inflectional classes for possession, and in particular that it presents a system of default marking for most third persons that cross-cuts most classes, as is expected from a system of inflectional classes. Data for this section is taken from van den Berg & Matsumura (2008). Jackson (2014) is also a recent survey of the language, with minor dialectal differences, but he does not describe the morphology of possession in as much detail as van den Berg & Matsumura (2008).

Alienables are the default class, used for most lexemes, including some body parts and most kin terms. Alienables are marked by the exclusive use of prefixal possessor marking on the possessee. The prefixes realize the values for number and person of the possessor, as indicated in Table 98:

	SG	PL
1INCL		it-
1EXCL	a-	am-
2	o-	e-
3	i-	ir-

Table 98. Irarutu default possessive marking (Jackson 2014:120)

So-called inalienables in fact cover a number of different patterns. The main pattern is for lexemes that add a suffix to the prefixal forms of alienably possessed nouns. This pattern is only available for first and second person, the third person marking being always only marked with a prefix. This is thus a case of default marking, common in inflectional systems. The opposite analysis would be to consider that the alienability distinction is neutralized in third persons, which would in any case be contrary to a semantic analysis of the alienability split as classes of lexemes reflecting some difference in meaning. The suffixes only mark person, not number. There are two classes of suffixes, for stems ending in a vowel and for stems ending in a consonant, as shown in Table 99:

	FREE PRONOUN	PREFIX	SUFFIX V-STEM	SUFFIX C-STEM
1SG	a ~ ja	a-	-g	-ûg
1PL.INCL	it	it-		
1PL.EXCL	am	am-		
2SG	o	o-	-m	-ûm
2PL	e	e-		
3SG	i	i-	-Ø	-Ø
3PL	ir	ir-		

Table 99. Circumfixal marking of inflectional class (1 and 2) in Irarutu

The two classes taking suffixes are not entirely constrained by phonology, as two words ending in a vowel still take the C-stem endings, *kki* 'chin' and *tî* 'liver. This is again showing that the inflectional class of a lexeme for the possessive paradigm of Irarutu is somewhat constrained or linked to extramorphological factors (here by phonology) but not determined by such factors.

Another subclass with three members has suffixes *-nûg* for 1SG ans *-nûm* for 2SG:

ROOT	MEANING	1SG	2SG	3SG
fgie	back	a-fgie-nûg	o-fgie-nûm	i-fgie
nmu	thigh	a-nmu-nûg a-nmu-g	o-nmu-nûm o-nmu-m	i-nmu
wo	spirit	a-wo-nûg	o-wo-nûm	i-wo

Table 100. Third pattern of circumfixal marking in Irarutu

In addition, one of the members of this subclass, *nmu* 'thigh', shows overabundance: first and second person marking can alternate between two different inflectional classes, yelding as a result two different forms for the realization of each cell, as shown in Table 100. The third person form is still default marked.

Another class is frozen compounds where only the first element of the compound is marked by a suffix, which gives an infixal like type of marking, as shown in Table 101:

ROOT	MEANING	1SG	2SG	3SG
frasi	elbow	a-fra<g>si	o-fra<m>si	i-frasi
ftarie	buttocks	a-fta<g>rie	o-fta<m>rie	i-ftarie
fasme	ankle	a-fa<g>sme	o-fa<m>sme	i-fasme
mtare	cheek	a-mta<g>re	o-mta<m>re	i-mtare

Table 101. First compounding pattern of inflection in Irarutu

Some compounds mark the suffix on both the first and the second element as shown in Table

102. This means that for that class of compounds, the exponence of possession is extended to three places of marking. In addition, for some of those frozen compounds, it is not possible to find equivalent words in the language that would make up the compound: the infixal marking is not transparently a suffixation in synchrony.

ROOT	MEANING	1SG	2SG	3SG
frafû	shoulder	a-fra<g>fû-g	o-fra<m>bû-m	i-frafû
mtarûm	face	a-mta<g>rû-g	o-mta<m>rû-m	i-mtarûm
wifu	knee	a-wi<g>fu-g	o-wi<m>bu-m	i-wifu

Table 102. Second compounding pattern of inflection in Irarutu

For the six classes described above, third persons use a default marking with prefixes alone. Such default marking is typical of inflectional class systems. For a number of classses, there will be a default marking of third persons: such persons are marked identically regardless of the inflectional class. The inflectional classes will thus only be distinctive for other person and number specification of possessor. They can in fact all be considered as subclasses in an inheritance hierarchy model.

Still another class is peculiar in marking both affixal elements as prefixes, and as having a distinctive second prefix *-n-* for third person, as shown in Table 103:

ROOT	MEANING	1SG	2SG	3SG
-fut	younger sibling	a-g-fut	o-m-but	i-n-fut
-îtn	child-in-law	a-g-îtn	o-m-îtn	i-n-îtn
-tamn	parent-in-law	a-g-tamn	o-m-tamn	i-n-tamn
-tat	grandchild	a-g-tat	o-m-tat	i-n-i-tat

Table 103. Double prefixing marking of possession in Irarutu

Finally, the remaining lexemes exhibit irregular marking (Table 104). This is yet another characteristic of inflectional class systems that is found in this so-called alienability split system: a number of lexemes are marked idiosyncratically, as if there were a number of classes with only one or two members. This is not expected if the system of classes were motivated semantically.

ROOT	MEANING	1SG	2SG	3SG
je	flesh	a-jo-g	o-jo-m	i-je
mmaje	tongue	a-majo-g	o-majo-m	i-mmaje
-rgun	head	a-rgun-ûg	o-rgun-ûm	i-rguin
tafad	older sibling	a-ta<g>fad	o-tabad	i-tafad
tantûf	friend, brother	a-ta<ge>ntûf	o-ta<me>ntûf	i-tantûf
wegur	nose	a-we<ge>gur	o-we<me>gur	i-wegur

Table 104. Irregular marking of possession in Irarutu

Irarutu shows a number of inflectional classes for what is usually the inalienable class. It cannot be simplified to two classes of alienables and inalienables, as neither of those classes is definable semantically. In particular, most kin terms and some body part terms inflect according to the larger, open class which would have to be called alienable in such a dual system. Irarutu is peculiar in showing a default marking for third persons for a number of those classes, which is typical of inflectional class systems. In addition, Irarutu only shows inflectional classes that are marked inflectionally on the head possessum noun, without the periphrastic class shown in other languages. It has developed into a clear system of inflectional classes for possessive paradigms.

7.4.4 Inflectional classes in possessive paradigms in other Austronesian languages

The type of inflectional class systems for possessive paradigms described until now for some South Halmahera-West New Guinea languages is also found in a number of other languages. Within Austronesian, I have found it on a number of languages spoken around New Guinea, in particular Ambel, a South Halmahera-West New Guinea language spoken in the Raja Ampat, in Wooi, a West Yapen language of South Halmahera-West New Guinea, and in Dobel, an Austronesian language spoken on the Aru archipelago. No doubt further examination of descriptive grammars would yield more examples.

7.4.4.1. Ambel

Ambel is particularly interesting in that it seems to have three classes that are at least partially correlated with semantics. In Ambel, there is a distinction between an alienable class marked

periphrastically, which is the type of realization for most nouns, an inalienable class marked inflectionally, and at least two irregular lexemes where possession is marked both affixally and through stem changes, with slightly different affixes (Laura Arnold p.c.). This is thus a minimal system of inflectional classes, compared to more complex systems such as Biak or Irarutu.

7.4.4.2. Wooi

Wooi is a South Halmahera-West New Guinea language spoken on the West of Yapen Island in Indonesian Papua (Sawaki 2016). The language distinguishes indirect from direct possession: indirect possession is marked periphrastically with a marker *ne* preposed to the possessed noun and taking agreement affixes agreeing with the possessor; direct possession is marked directly on the possessed head noun (Sawaki 2016:127-132). Lexemes presenting head-marking of possession include body parts, kin terms, some bodily products, and some abstract nouns. Singular forms realize possession through suffixation, non-singular forms through extended exponence marked by both prefixes and suffixes. It is possible to further distinguish subclasses in direct possession, which mostly affect singular forms, distinguishing body parts from kin terms and compound nouns, in addition to the indirect construction (Table 105). One should note that these subclasses are not definable through semantics only: the class of compounds is a morphological one, and at least two body part terms, *huhu* 'breasts' and *rerawa* 'skin', inflect according to the indirect construction. Table 105 shows the inflectional classes distinguished for possession in Wooi (Sawaki 2016:159) for singular possessor forms only:

	Direct possession			Indirect possession
	Simple body parts	Compounds	Kinship	Common nouns
1	Ø	Ø	Ø	Ø
2	-mu	-ng	-mu	-mu
3	-ng	Ø	-n	-i

Table 105. Inflectional classes in singular possessive paradigms in Wooi (Sawaki 2016:159)

An interesting point for inflectional classes in Wooi is that similar exponents can mark different feature value combinations: thus *–ng* can mark either 2SG possessor on compounds or 3SG

possessor on simple body part terms. It is also a case of default marking, as the three classes of so called inalienables share identical realizations for all nonsingular possessor marking.

7.4.4.3. Dobel

In Dobel (Hughes 2000), there is a contrast between an alienable class marked periphrastically, and a number of inalienable classes marked inflectionally through possessor indexing affixes. Inalienables include most kinship terms and body parts, but not all of them, as well as some locational nouns. The inalienable classes are thus not definable precisely through semantics, but correspond to inflectional classes for possessive paradigms. Dobel is interesting in that it presents a number of default class markings, with subclasses in an inheritance hierarchy model type.

The alienable paradigm is shown in Table 106. These forms can function as possession markers with a possessed noun, or as full pronouns, in which case they function as non-active verbs and take undergoer suffixes agreeing with the absent possessed noun (Hughes 2000:146). They can sometimes be accompanied by a personal pronoun form, in particular to disambiguate homophonous forms of 1PL.INCL and 3PL.

	SG	PL
1INCL		ʔada
1EXCL	ʔana	ʔama
2	ʔamu	ʔami
3	ʔani	ʔada

Table 106. Alienable possessive markers in Dobel (Hughes 2000:146)

For a restricted number of lexemes, the marking of possession is realized through affixation. The main paradigm is shown in Table 107:

	SG	PL
1INCL		-da
1EXCL	-ŋu	-ma
2	-m	-mi
3	last vowel of root -> *i*	-di

Table 107. Affixal marking of possession in Dobel: first pattern

Roots ending in *-n* or in *-r* have 2SG zero marked, a subclass that could still be accounted for

phonologically. On the other hand, a number of lexemes whose stem ends in *-a* take two different subclasses of marking: for some of them, the last vowel of the stem is replaced by *-i* (for example *sama* 'father', Table 108), for others, a yod is suffixed to the stem (for example *yaba* 'leg', Table 109). This pattern of alternation cannot be accounted for by morphophonemic changes. These two classes show a default marking of most cells, and only one cell is marked specifically for this class, which is typical of inflectional class systems.

	SG	PL
1INCL	sama-ŋu	sama-da
1EXCL		sama-ma
2	sama-m	sama-mi
3	sami	sama-di

Table 108. Inflection of sama 'father' in Dobel (Hughes 2000:144)

	SG	PL
1INCL	yaba-ŋu	yaba-da
1EXCL		yaba-ma
2	yaba-m	yaba-mi
3	yaba-y	yaba-di

Table 109. Inflection of yaba 'leg' in Dobel (Hughes 2000:144)

A number of other lexemes show metaphony of the stem vowel for 3SG. This vowel change can be accounted for in diachrony through metaphony triggered by a suffix *-i*, but it cannot be accounted for phonologically in synchrony (Table 110). Finally, one lexeme shows suppletion for its 3SG cell.

	SG	PL
1INCL	lar-ŋu	lar-da
1EXCL		lar-ma
2	lar-m	lar-mi
3	ler	lar-di

Table 110. Inflection of lar *'voice' in Dobel (Hughes 2000:145)*

Another group of lexemes, still making use of default marking for most of the paradigm, has specific 2SG marking, as shown in Table 111. It can be considered as a subclass of the 3SG yod marking class. These also seem to show overabundance for that same cell, having two possible

realizations:

	SG	PL
1INCL		yana-da
1EXCL	yana-ŋu	yana-ma
2	yana-w yana-ʔu	yana-mi
3	yana-y	yana-di

Table 111. Inflectional of yana *'child' in Dobel (Hughes 2000:145)*

Finally, locational nouns and a number of other nouns are defective, just as they are in Biak. Here they only present third persons.

7.4.5 Interim conclusion

Inflectional classes in possessive paradigms are widely found in South Halmahera-West New Guinea languages, and also present in other Austronesian languages, such as Dobel. They always take a rather similar form: multiple classes arise for those lexical items that would typically be classified as inalienable. The typically alienable class is not further divided. One should note that in many Oceanic languages the so called alienable class is commonly further divided into multiple classes, but these classes have a clear semantic basis. In Sawai for example, a South Halmahera-West New Guinea language, edible items are opposed to generally possessed items (van den Berg 2009). The assignment of nouns into such classes is purely semantic, their expression periphrastic, and they thus fall outside the domain of enquiry of this chapter. Such systems are often said to allow recategorization, but recent works show that they are closer to gender systems, particularly because recategorization is only limited to a few items, such as 'coconut' (Franjieh 2016).

There is some semantic component to the creation of inflectional classes in possessive paradigms in Austronesian languages in that they typically include at least some body parts and kin terms. But those classes cannot be defined in terms of semantic assignment of lexemes to classes, because none of the languages studied in this section assign all body parts or kin terms to such classes, but always assign some such terms to the alienable class.

Heteroclisis is a characteristic of inflectional class systems, and if a language presents a mixture

of classes of inflection, it proves that it presents inflectional classes. Some languages show heteroclisis in their system of inflectional classes for possessive paradigms, in particular between what has traditionally been termed the alienable and the inalienable classes, which confirms that those are inflectional classes. This is in particular the case in Biak and in Ambai.

Other languages present default marking for parts of the paradigm, and an inheritance hierearchy like model of classes. Such usual organization of inflectional classes is present in Dobel and in Irarutu. Ambel divides its classes on a semantic basis that might explain how such classes initially developed from semantically motivated class assignment, before some blurring occurred and those semantic distinctions became bleached. Biak still shows traces of such initial conditioning.

7.5. Possessive classification as inflectional class: evidence from non-Austronesian languages from New Guinea

A number of non-Austronesian languages from New Guinea present inflectional classes for their possessive paradigms. While some languages present systems of classes that are quite comparable to the Austronesian languages of the area, as is the case in Amele, a number of languages present rather different systems, which probably have different historical origins.

7.5.1 Amele

Amele is a Nuclear Trans New Guinea language spoken in Papua New Guinea (Roberts 1987; Hammarström et al. 2017). It is in particular characterized as having a very large number of classes for possessive paradigms, amounting to 32 classes (Roberts 1987:171-175). Most lexemes realize possession through a genitive marker *na* which functions as a postposition (Roberts 1987:171):

16)

Ija	na	jo	mane-i-a
1SG	of	house	burn-3SG-TOD.PAST

'My house burned down' (Roberts 1987:171)

The remaining lexemes bear inflection for possessor agreement realized through a suffix. These include lexemes whose meaning is typical of an inalienable class, mainly kin terms and body parts, but also a number of abstract nouns such as *gelehi* 'bravery', *cehewa* 'wealth' or *oloho* 'bad temper'. There is a further division between kin terms and other lexemes: kin terms additionnally encode the number of the possessee, which gives two rather different paradigms. This distinction cross-cuts the division into classes, although class 25 to 31 mostly include kinship terms, with also some body part terms. The main paradigm type is given in Table 112 for *dewe-g* 'his body' (class 1, underlined vowels are epenthetic harmonic vowels; Roberts 1987:171). For both paradigm types, there is systematic syncretism between second and third dual possessor forms, and between second and third plural possessor forms.

	SG	DU	PL
1	dewe-ni	dewe-ni-le	dewe-ni-ge
2	dewe-n	dewe-n̲e-la	dewe-n̲e-ga
3	dewe-g	dewe-n̲e-la	dewe-n̲e-ga

Table 112. Inflection of dewe *'body', class 1, in Amele (Roberts 1987:171)*

The paradigm for kin terms presents the same systematic syncretisms, and simply adds forms for plural possessee, as shown in Table 113 for the lexeme *cot-ig* 'his brother' (class 27, underlined vowels are epenthetic harmonic vowels; Roberts 1987:171-172):

	SG	DU	PL
1 / SG possessee	cot-i	cot-i-le	cot-i-ge
1 / PL possessee	cot-i-el	cot-i-le-il	cot-i-ge-il
2 / SG possessee	cot-in	cot-o̲-la	cot-o̲-ga
2 / PL possessee	cot-in-el	cot-o̲-la-il	cot-o̲-ga-il
3 / SG possessee	cot-ig	cot-o̲-la	cot-o̲-ga
3 / PL possessee	cot-ug-ul	cot-o̲-la-il	cot-o̲-ga-il

Table 113. Inflection of cot *'brother', class 27, in Amele (Roberts 1987:171-172)*

There are 109 lexemes that take possessive suffixes, divided into 31 different inflectional classes. Those classes are given in Table 114 (data from Roberts 1987:172-175), distinguishing forms of the singular, together with the number of lexemes in each class. It should be apparent that

a number of classes in fact correspond to idiosyncratic realizations for only one lexeme, while other classes have up to fifteen members.

Class number	1SG	2SG	3SG	Number of lexemes in class
1	-ni	-n	-g	14
2	-ni	-nin	-g	4
3	-ni	-in	-g	5
4	-ni	-n	-ug	1
5	-eni	-ein	-ug	1
6	-ni	-n	-nag	1
7	-ni	-n	-nug	1
8	-mi	-m	-g	2
9	-ni	-nin	-n	4
10	-ni	-in	-n	15
11	-ni	-n	-c	3
12	-ni	-nin	-c	2
13	-ni	-in	-c	3
14	-ni	-in	-ic	1
15	-mi	-m	-c	1
16	-mi	-im	-c	2
17	-mi	-m	-h	11
18	-mi	-im	-h	2
19	-ni	-n	Ø	4
20	-ni	-in	Ø	1
21	-ini	-in	Ø	4
22	-ini	-inin	Ø	1
23	-ani	-ain	Ø	1
24	-mi	-m	Ø	7
25	-ni	-in	-ig	1
26	-li	-in	-ig	1
27	-i	-in	-ig	7
28	-i	-im	-ig	1
29	-i	-in	-iag	2
30	-i	-in	-ag	4
31	-i	-en	-eg	2

Table 114. Possessive inflectional classes in Amele (after Roberts 1987:172-175)

It should be evident from the forms given in Table 114 that a lot of those classes have much in common. A number of the classes make use of defaults such as 1SG *-ni* or *-i*. The presence of multiple default forms shows that Amele also countervenes the No-Blur Principle. Still, it is not possible to find any clear phonological conditioning for the variants. For example, thirteen classes include nouns that end in the vowel *-a*, as shown in Table 115 (after Roberts 1987:172-175; only

one lexeme is given for each class).

CLASS	ROOT	MEANING
1	taba	body (with head)
3	maja	shame
4	cula	pride
8	cebina	man's sister
10	hulufa	scar
13	gola	blood
14	gada	hips
16	sahala	saliva
17	mela	son
18	beila	tongue
20	ola	face
25	jaja	great great grand parent/child
26	tana	wife's father

Table 115. Absence of phonological conditioning for inflectional classes in Amele: -a stem nouns (after Roberts 1987:172-175)

Similarly, closely similar roots belong to different inflectional classes. This is for example the case of *men* 'brother's wife / husband's sister' which belongs to class 27, and *min* 'vagina' which belongs to class 30. The markers are also not limited to a monovalent realization of features. Final *-n* for example is used to realize 3SG for classes 9/10, but it realizes 2SG for a number of other classes. Both cases include some of the largest classes: *-n* realizes 2SG in class 1, which comprises 14 lexemes, and 3SG in class 10 which comprises 15 lexemes.

It is highly possible that the different classes arose from initially motivated phonological adjustments. There are also traces of some markers being initially more restricted semantically to only a subclass. The initial situation may have been close to that observed now in Ambel, with markers used on the one side for body parts, on the other for kin terms. In Amele, "all nonmarriage nuclear family terms are marked with *-i* underlyingly for first person singular" (Roberts 1987:172). One could thus hypothesize that the marker *-i* was initially restricted to nuclear family terms, and was later bleached and extended to other terms such as the words for 'eye' or 'vagina'.

The system of inflectional classes of Amele, although more developed, closely resembles the

systems analysed in the previous section for Austronesian languages: it is restricted to lexemes that are commonly assigned to inalienables, and it makes extensive use of default marking. It is different in that some markers are polysemous.

7.5.2 Anêm

Anêm, a language isolate spoken in the West New Britain province of Papua New Guinea, presents a large number of inflectional classes for its possessive paradigm (Thurston 1982). Anêm distinguishes 21 classes, plus at least four irregular paradigms. Those classes are interesting in that they can easily be represented in an inheritance hierarchy model. Anêm distinguishes four major classes depending on the way possessor agreement is realized for the features person, number and gender of the possessor, as shown in Table 116. There are two minor variants of these classes where 2SG is marked differently, with a marker *-r*. There is systematic syncretism between 2SG and 3PL for class 1 and 2 only, but not with the *-r* variant. Class 3 and 4 show gender syncretism for 3SG. Class 4 also shows systematic syncretism of third persons, and a number of variants that are correlated with stem changes (see below), as well as overabundance for the same cells.

	1	2	3	4
1SG	-i	-e	-a	-at
2SG	-î	-ê	-îr	-ir
3SG.MASC	-u	-o	-î	-it / -il
3SG.FEM	-îm	-êm	-î	-it / -il
1PL.INCL	-iŋ	-eŋ	-iŋ	-nis
1PL.EXCL	-în	-ên	-în	-nit
2PL	-îŋ	-êŋ	-îŋ	-ŋît
3PL	-î	-ê	-i	-it / -il

Table 116. Macro-classes for possessive paradigms in Anêm (after Thruston 1982:87-89)

All the other classes are subclasses of the four main classes presented in Table 116. This means that the language presents a system of default affixal marking. Other classes are formed with a stem augment followed by the possessor agreement marker or one of its variants. For example *nan-* 'garden' is of class 1 but is realized with the augment *-u-* on the stem. The class markers do not seem to be correlated with semantics, nor does it seem that there is one massively productive class as opposed to all others. Class 1 for example includes kin terms such as father, body parts such as knee

or eye, but also objects such as lime powder or concepts such as garden. The system is thus quite different from the Austronesian languages in the previous section, or even Amele.

It is possible to represent the system of classes in Anêm in an inheritance hierarchy model. Figure 1 is such a representation:

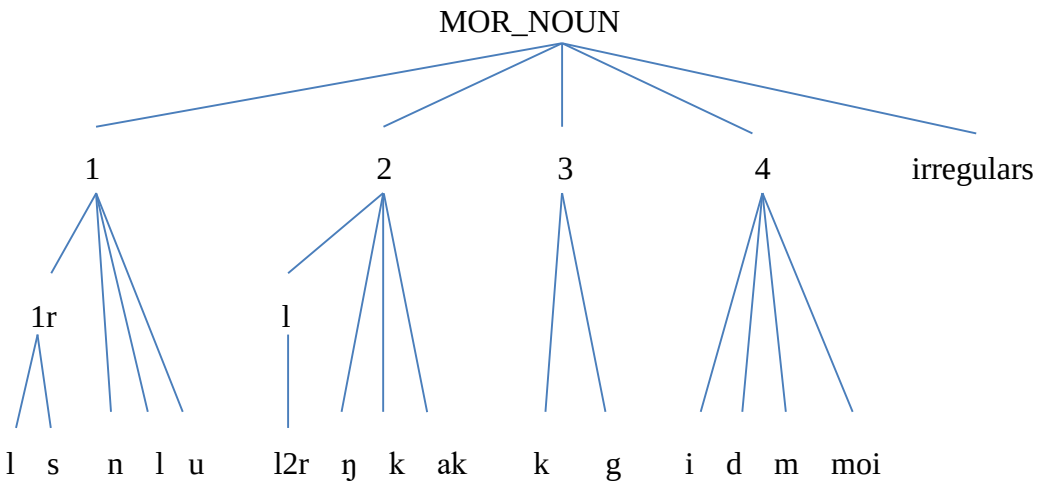


Figure 3. Inheritance hierarchy for noun morphology in Anêm

The system does not seem to work in the same way as languages which originally first distinguished alienables from inalienables. Here, there is no specific group of classes for typically inalienable lexemes. It is still possible to imagine a scenario for the development of the markers. Agreement markers appear to be partially related to one another. For example class 1 to 3 plural markers are clearly variants of each other, class 1 and 3 making use of some default marking. One could posit that such markers evolved from pronouns that were used to index the possessor on a possessive construction, which then gradually cliticized and finally morphologized. The differences in form could in part be due to the augment element. As to the augments, the only ready explanation would be that they originated as possessive classifiers, of the kind now common in Oceanic languages: there is a range of classifiers indexed with the possessor, and the choice depends on the lexical semantics of the possessed noun. Thurston (1982) only gives one example of lexeme per class, which means that it is not possible to check such a hypothesis. The fact that Anêm is a language isolate also prevents any comparative attempt.

7.5.3 Salt-Yui

Salt-Yui is a nuclear Trans New Guinea language (Irwin 1974; Hammarström et al. 2017). It generally marks possession through the use of a pronoun preposed to the possessed noun:

17) na heba dungwi

my sweet.potato it.is

'It is my sweet potato' (Irwin 1974:29)

Inalienables distinguish the marking of kin terms and of body parts. Kin terms are marked by a suffix and an optional kin marker *-bi*, as shown in Table 117 (Irwin 1974:28):

1SG	eu-na-bi / eu-na
2SG	eu-ni-bi / eu-ni
3SG	eu-ng-bi/ eu-ng

Table 117. Kinship terms possessive paradigm in Salt-Yui: eu- *'wife' (Irwin 1974:28)*

In addition, two kinship terms are irregular, showing both prefixes and suffixes, in addition to stem suppletion. These are the words for mother and father.

There is another slightly different class for body parts and a number of what Irwin calls property nouns, including *oo* 'house'. The class is thus not definable through semantics only, although it seems from the description that it includes all body parts. The difference is shown in Table 118 (Irwin 1974:29):

	Kin terms	Body parts
1SG	-na-(bi)	-na
2SG	-ni-(bi)	-n
3SG	-ng-(bi)	-ng

Table 118. Kin terms and body parts possessive marking in Salt-Yui (Irwin 1974:29)

If one excepts the optional marker *-bi* for kin terms, the markings are quite similar and show a system of default marking. Only 2SG is different. Salt-Yui thus presents a minimal system of inflectional classes, probably in the initial stages of its development. There are two classes marking possession affixally, which happen to be minimally differentiated with markers that have a clear common origin. Those classes are highly correlated with semantics. The remainder of nouns are

marked by a periphrastic construction.

7.6. Possessive classification as inflectional class: evidence from other language families

A number of languages also show inflectional classes for their possessive paradigms outside the New Guinea area. Of those, this section particularly centres on Eastern Pomo, Maricopa, and Ungarinjin.

7.6.1 Eastern Pomo

Eastern Pomo is a Pomoan language of California (McLendon 1975). It has been described as presenting a basic contrast between alienable and inalienable possession, which in fact covers a distinction between periphrastic, pronominal marking and head-marking. The alienable markers are in fact the genitive form of the personal pronoun, agreeing in number, gender and person with the possessor, as shown in Table 119 (after McLendon 1975:113, only full forms are shown; the dot indicates vowel length):

	SG	PL
1	wáx	wáybax
2	míbax	maybax
3.MASC	mí·pibax	bé·kibax
3.FEM	mí·ribax	

Table 119. Alienable possessive marker in Eastern Pomo (McLendon 1975:113)

Kin terms are head-marked for possession, and take one of two possible sets of prefixes. There are thus two inflectional classes for kin terms, as shown in Table 120. The second set of prefixes seems to show a marker *-ma·-* which is also the anaphoric third singular pronoun restricted to that class. It also shows syncretism for all third person forms. Finally, the system also shows default marking for 2SG. The link between the possessive markers in the head-marking constructions and the genitive pronouns is quite evident. One can thus postulate a grammaticalization path that targets

terms for special phonological reduction of the markers, maybe for reasons of frequency.

	SET 1	SET 2
1.SG	wí-	wíma·-
2.SG	mi·-	mi·-
3.SG.MASC	mí·pi-	há·mi·-
3.SG.FEM	mí·ri-	há·mi·-
1.PL	wáy-	wáyma·-
2.PL	máy-	máymi·-
3.PL	be·ki-	há·mi·-

Table 120. Possessive prefixes in Eastern Pomo (McLendon 1975:113)

Set 1 markers realize possession on only three kinship terms, the words for wife, husband, and second or third person's child. Set 2 markers occur with twenty-two kinship terms (McLendon 1975:114).

7.6.2 Maricopa

Maricopa is a Cochimi-Yuman language spoken in Arizona (Gordon 1986; Hammarström et al. 2017). It can be described as having a basic tripartition of possessives between inalienables, other possessable items, and alienable items (Gordon 1986:30-35). Inalienables are also obligatorily possessed. Most body part and kin terms, as well as clothing items, belong to the class of inalienables. Possession of those items is realized through a set of number neutral prefixes, as shown in Table 121:

	'hand'	'younger brother'
1	'-iishaaly	'-shchaa
2	m-iishaaly	m-shchaa
3	iishaaly	shchaa

Table 121. Maricopa inalienable possessives (Gordon 1986:30)

There is a second class of lexemes with head-marking of possession, for which a prefixal support *ny-* must be affixed to the stem before inalienable affixes are affixed. Some kin terms, such as *kwr'ak* 'husband', belong to that class (Gordon 1986:32). There are also some nouns that vary between the two classes, showing overabundance, such as *shlymak* 'back'. This shows that the two classes are not entirely definable through semantics.

Finally, the main class of possession makes use of a periphrastic expression, postposing *nywish* or *(ny)uuwish* with possessor inflection to the possessed noun. In the case of domesticated animals, the language further makes use of a classifier *ny-hat* (originally 'POSS-dog') postposed to the possessed item (Gordon 1986:33).

7.6.3 Ungarinjin

Ungarinjin or Ngarinyin is a Worrorran language from the Kimberley region in Australia (Rumsey 1982; Hammarström et al. 2017). It seems at first sight to distinguish three classes for possession, based on meaning: prefixing for body parts, suffixing for kinship terms, and periphrastic for the remainder of nouns. It could be said to distinguish alienable from inalienable possession on the basis that typically inalienable items show head-marking. On closer inspection, the system is more complex than just three classes perfectly defined by semantics: some body parts are prefixing, others are inflected according to the alienable pattern with a genitive pronoun. There seem to be an equal number of terms in each class, and no subcriterion can be found to group them semantically (Rumsey 1982:46).

For body part terms, there seem to be some phonological adaptation of the affixes, which could be considered as subclasses. This is in particular the case for the third singular of gender W cell. Some phonological adaptation can be seen for plural forms of *amuḷar* 'forehead', which are not based on the 3SG stem as is the case for the other two examples given here. The same lexeme shows overabundance for 3PL. Table 122 gives the example of three body part paradigms (Rumsey 1982:43):

	‘bone(s)’	‘foot (feet)’	‘forehead’
1SG	ŋiyoṇar	ŋiyembularu	ŋiyamuḷar
2SG	njuŋoṇar	njuŋembularu	njuŋamuḷar
3SG.MASC	oṇar	embularu	amuḷar
3SG.FEM	njoṇar	njembularu	njamuḷar
3SG.M	moṇar	membularu	mamuḷar
3SG.W	wuṇar	wembularu	wumuḷar
1PL.INCL	ŋaroṇar	ŋarembularu	ŋarumuḷar
1PL.EXCL	njaroṇar	njarembularu	njarumuḷar
2PL	guroṇar	gurembularu	gurumuḷar
3PL	buroṇar	burembularu	burumuḷar / buramuḷar

Table 122. Body part possessive paradigms in Ungarinjin (Rumsey 1982:43)

Kin terms realize possession through suffixes that encode person and number of the possessor, and number of the possessed, for which it only distinguishes two values, singular and plural, where the language generally distinguishes a dual as well. In addition, 3SG does not distinguish gender. The main realizations are given in Table 123 (Rumsey 1982:47):

POSSESSOR	POSSESSED	
	Singular	Plural
1SG	-ŋi	-ŋiri
2SG	-ni	-niri
3SG	-naŋga	-naŋgari
1PL.INCL	-ŋaruna	-ŋarumbu
1PL.EXCL	-njaruna	-njarumbu
2PL	-nudna	-nurumbu
3PL	-yiduga	-(wur)umburu

Table 123. Kin terms marking of possession in Ungarinjin (Rumsey 1982:47)

In addition, there is a subclass of nouns for 2SG cell: for some nouns, the suffix changes a final *-a* stem word to *-i* for assimilating stems, and does not for non-assimilating stems. The distinction is lexical. Thus *mara* ‘wife’ is assimilating, while *maḷa* ‘daughter (of a woman)’ is not. Finally, two lexemes are irregular, showing suppletion for first person: *ira* ‘father, son’, and *ŋara* ‘mother’. Ungarinjin thus shows inflectional classes for possession which are partially conditioned on the semantics of the lexemes. In addition not all lexemes are marked for the same values, which should in theory prevent one from analysing these classes as inflectional classes.

Worrorra, from the same language family also distinguishes two different classes of inalienables, body parts where possession is realized through prefixes, and kinship terms where possession is

realized through suffixes (Clendon 2014). Alienables are marked with a genitive/dative pronoun. The distinction should thus probably be reconstructed for proto-Worrorran, where it may have originated as a semantics based system of classification of possession opposing body parts, kinship terms and the remainder of nouns. It is not possible to assert the directionality of the evolution of the system, but the fact that not all inalienables use the same type of affixation (prefixes for body parts and suffixes for kin terms) advocates for successive stages in the morphologization of the distinctions. This is further proven by the fact that the two head-marking constructions do not distinguish the same set of feature value pairings: kinship terms also encode number with reference to the possessed item, in addition to person and number of the possessor; body parts do not encode the number of the possessed head noun.

7.6.4 Other languages

Inflectional classes in possessive paradigms are also attested in Slave (Rice 1989), Burushaski (Nichols and Bickel 2013) and in the Zamucoan languages old Zamuco, Ayoreo and Chamacoco (Ciucci & Bertinetto 2017).

7.7. Conclusion

A large number of languages from all areas and various different language families exhibit inflectional classes for their possessive paradigm. The phenomenon is present at least in Australia, in and around New Guinea, and in the Americas. A number of language families exhibit this phenomenon, including Austronesian, Trans New Guinea, Worroran, Pomoan, Cochimi-Yuman, Zamucoan, and the language isolates Anêm and Burushaski.

All those languages present head-marking of the possessive relationship for at least inalienables: possession is encoded morphologically on the possessed head noun through a set of affixal markers which cross-reference the possessor. Such head-marking seems to be a prerequisite for the development of inflectional classes specifically in possessive paradigms.

Most languages (except Anêm and Burushaski) have similar forms for their inflectional classes. Whatever the exact number of classes, going from two to 33, the languages presented in this chapter exhibit a set of small, closed-membership classes which include lexical items that are typically considered inalienable, mostly body parts and kin terms, opposed to a larger, productive inflectional class, most often marking possession periphrastically. None of the languages investigated presents inflectional classes that are perfectly amenable to a semantic definition. Some, but not all, kin terms and body parts belong to each class, sometimes with other types of lexemes, without it being possible to define precisely each class in semantic terms.

The fact that these languages present relatively similar systems does not mean that such class systems all evolved in the same way. Thus in Eastern Pomo the system evolved with a differential phonological reduction of pronominal possessive markers for those lexical items that are most commonly possessed, that is a subset of kin terms. In other languages, such as Biak, one can reconstruct the appearance of a new construction (the so-called alienable), which was blocked by pre-existing constructions on those items that are most commonly possessed, in that case a subset of body parts and kin terms. Both types of possible evolution mentioned by Haspelmath (2017) are thus present in the sample languages. The evolution of inflectional classes in possessive paradigms is thus tied in with sound change on one side, and grammaticalization on the other. But these happen in a very specific type of environment, where classes may have already been distinguished through lexical semantics.

Anêm and Burushaski are different in that they do not present a general, productive class opposed to smaller, closed membership classes. In Anêm for example, classes seem to be evenly distributed across the lexicon. This probably indicates that these classes evolved from different systems. For Anêm at least, it is possible to hypothesize an origin in possessive classifiers.

Alienability distinctions have not been previously considered as inflectional classes. This chapter has shown that they should be considered so in a large number of languages. This has implications for the way we treat inflectional classes in morphological theory. One the one hand, it

is fairly possible that a number of distinctions that have been made about particular systems on the basis of a rather shaky analysis based on a semantic feature have to be reconsidered as purely morphological phenomena. Similar alternations may arise in other contexts, as will be shown in the next chapter for gender systems. On the other hand, developments which make use of the common mechanisms of grammaticalization or sound change, when applied to specific environments which are tied to some semantic conditioning, seem to give rise to rather similar systems of classes, ones where there is a large open class, often marked periphrastically, and a number of closed, small membership classes, marked affixally.

8. Inflectional classes from agreement classes

The preceding chapter has shown that systems of inflectional classes can originate in systems of alienability distinctions in possessive paradigms. In such a case, the system of inflectional classes does not arise through a specific process of reanalysis or grammaticalization, but through the reanalysis of an entire, preexisting system, which becomes a system of inflectional classes, through processes of sound change and grammaticalization applied to that existing system. In a similar way, this chapter is concerned with the transformation of a preexisting system that partitions the lexicon into classes, which becomes a system of inflectional classes. This system is the gender system exhibited by a range of languages, which, when marked on nouns, can give rise to a system of inflectional classes. This potential source of inflectional classes has received little attention in the literature.

Gender is the division of nouns into multiple classes "as reflected in the behavior of other words" (Hockett 1958). Therefore, gender is characteristically shown through agreement classes: each noun controls the agreement of one or multiple targets according to the gender it belongs to. Thus, in Examples 1 and 2 from French, the nouns *garçon* 'boy' and *fille* 'girl' control agreement on the definite article and on the adjective, according to their gender classification, masculine in Example 1, feminine in Example 2:

1) le petit garçon
   the.M small.M boy
   'the small boy'

2) la petite fille
   the.F small.F girl
   'the small girl'

Gender is thus different from inflectional class, which is another form of classification dividing a word class into multiple classes, in that inflectional classes are marked on nouns, whereas gender

is shown through agreement on targets. In some languages, gender can also be realized on nouns, in addition to the agreement realization. In such cases one speaks of overt gender, because the system of classification is overtly marked on nouns, as opposed to covert gender systems. In overt gender systems, there can be some interaction between gender as overtly realized on nouns, and inflectional classes, which also show on nouns. Typically in such cases, the inflectional class concerns the marking of number.[49] Such systems are relatively common in Niger-Congo languages (see Section 8.3.), but can also be shown to largely operate in more familiar languages such as Spanish or Italian (see Section 8.1).

Contrary to most systems of inflectional classes, the assignment of lexemes to specific genders tends to follow systematic rules based on either the lexical semantics of the noun, or its form. Some systems operate in a purely semantic fashion. Others use a mixture of semantic and formal assignment, based on the animacy hierarchy: nouns belonging to higher parts of the hierarchy are assigned semantically, while nouns below a certain threshold are assigned a gender based on their form. This is what is found in French for example, where higher animates (humans and some animals) receive semantic assignment to feminine or masculine gender based on the sex of the referent. Some animals and inanimates receive a purely conventional gender assignment, which can be shown for most nouns to be linked to the phonological shape of the word ending. Finally, a number of nouns form the residue, which is assigned lexically (Corbett 1991).

Greenberg (1978) sheds light on the possible origins of overt gender marking on nouns as arising from a system of gender marking through agreement. He states that this marking can originate in the fusion of an article marked for gender, which linearly appears close to the noun. There are four stages in the development of definite articles, which are the steps in the grammaticalization of such articles into overt gender markers. First they are definite articles, which then have their functional range extended to cover all non-generic but specific denotations. In the final stage, the article is extended to all nouns, even in generic uses, and becomes a marker of

49 This is probably due to the fact that number is one of the most prominent features marked on nouns crosslinguistically.

nounness. If the article was marked for gender, when it finally fuses morphologically with the noun, it becomes an affix marking gender directly on the noun. Now if the article also marked another inflectional feature, typically number, it means that what is attached to nouns is also a marker of number. The alternations of gender and number marking on the noun are thus now inflectional classes which also denote unambiguously the gender of the noun. Greenberg gives examples of languages at each stage of development. This is an example of an implicational relationship between gender and inflectional class, where the gender system also historically gave rise to the inflectional class distinction through grammaticalization of an article into an affix marking gender and another inflectional feature. Inflectional classes in such systems thus arise from the combination of further grammaticalization and an already existing system of classes determined by the feature gender. The newly created inflectional classes will thus show a very strong implicative relationship between the classes distinguished in gender, and the classes distinguished by inflectional classes.

A large number of languages show implicational relations between the gender of a noun and its inflectional class. It is thus highly possible that at least in some of those languages, a system of inflectional classes may have arisen from a gender system marked on nouns. In this chapter, I first give an overview of implicational relations between gender and inflectional classes in languages of Europe that have featured widely in the literature (Russian, Italian, Spanish). I then examine the relationship between gender and inflectional classes in some languages of New Guinea, following pioneering studies by Aronoff (1994) on Yimas and Arapesh: I analyse Bukiyip in comparison to Arapesh, to show that the system of inflectional classes probably arose from a system of gender marking. I then compare such systems of marking with the gender system of Miraña, which can shed some light on how a system of inflectional classes can develop from a system of gender with alliterative concord. Finally, I examine how systems of inflectional classes arose from a gender system in Niger-Congo languages, focusing particularly on the Gur languages.

8.1. The implicative relationship between gender and inflectional

class

In many languages there is an implicational link between the inflectional classes of nouns and their gender. This does not mean that there is a common origin for the two systems, but that they are very closely linked. For example, the rules for gender assignment may be sensitive to the inflectional class a noun belongs to: all nouns in a given inflectional class will have the same gender (a situation exemplified by Russian, see section 8.1.2.). Conversely, all nouns of a given gender could be part of the same inflectional class, in which case the predictability would go in both directions: gender predicts inflectional class perfectly, and inflectional class predicts gender perfectly. This ideal situation is not expected to be found, though, in languages, and there are mostly unidirectional implications. This section examines the relationship between gender and inflectional class in Spanish and in Russian, two languages which have featured widely in the literature on this topic.

8.1.1 Spanish

Spanish nouns roughly fall into three different inflectional classes: nouns suffixed in *-a*, nouns suffixed in *-o*, and suffixless nouns (Aronoff 1994:67). Suffixless nouns belong to two different subclasses, those for which there is insertion of an epenthetic vowel *-e*, and those which end in a consonant (Aronoff 1994). There is a rough implicational link between those inflectional classes and gender: most nouns in *-o* are masculine, although there is a handful of nouns such as *mano* 'hand' which are feminine; nouns in *-a* are mostly feminine, although there are about 600 nouns in *-a* which are masculine, most of which are characterized by having a feature +ANIMATE,[50] and most of which are names of professions and employments; finally nouns in the third class can be of either gender. The implicational relationship is thus far from ideal, but it works as a statistical tool for predicting the gender of nouns in the first two classes. Table 124 summarizes the situation:

50 A notable exception is *día* 'day'.

Marker	Class	Gender	Example	Gloss
-o	1	masculine	muchacho	boy
-o	1	feminine	mano	hand
-a	2	masculine	día	day
-a	2	feminine	muchacha	girl
-Ø	3	masculine	Cid	Cid
-Ø	3	feminine	sed	thirst
-Ø (*e* inserted)	3	masculine	padre	father
-Ø (*e* inserted)	3	feminine	madre	mother

Table 124. Inflectional class and gender in Spanish (Aronoff 1994:67)

Obviously, some other cues are useful for predicting the gender of nouns ending in a consonant in SPanish, such as morphological cues: some derivational affixes will always take the same gender, as is the case for abstract nouns in *-cion* which are always feminine.

8.1.2 Russian

In Russian, the implicative link between gender and inflectional class concerns the assignment of nouns to a specific gender: inflectional class is a good predictor of the gender a given noun belongs to. Russian thus presents a morphological system of gender assignment (Corbett 1991). The account presented here is mostly based on Corbett (1991:34-43).

Russian has three genders, masculine, feminine, and neuter. Nouns in Russian are assigned to genders following first a semantic rule of assignment for nouns denoting humans and higher animates: sex differentiable nouns denoting males are assigned the masculine gender, those denoting females are assigned the feminine gender. But this rule only covers a small fraction of Russian nouns. For other nouns, there is a rule of morphological assignment sensitive to inflectional class: the inflectional class a noun belongs to implicates its gender (this rule works for all but about 20 irregular lexemes). Table 125 presents the four main paradigms of inflectional classes for nouns in Russian. Nouns belonging to class I are masculine; those belonging to class II and III are feminine; those in class IV are neuter.

	I	II	III	IV
SINGULAR	'law'	'school'	'bone'	'wine'
Nominative	zakon	škola	kost'	vino
Accusative	zakon	školu	kost'	vino
Genitive	zakona	školy	kosti	vina
Dative	zakonu	škole	kosti	vinu
Instrumental	zakonom	školoj	kost'ju	vinom
Locative	zakone	škole	kosti	vine
PLURAL				
Nominative	zakony	školy	kosti	vina
Accusative	zakony	školy	kosti	vina
Genitive	zakonov	škol	kostej	vin
Dative	zakonam	školam	kostjam	vinam
Instrumental	zakonami	školami	kostjami	vinami
Locative	zakonax	školax	kostjax	vinax

Table 125. Inflectional classes in Russian nouns (Corbett 1991:36)

The implicative relationship between inflectional class and gender in Russian is shown further when nouns change inflectional class: in such cases, they also change gender (Corbett 1991:98-99).

The implicative link between inflectional class and gender is by no means systematic in all languages. For example, in French, gender assignment mostly depends on the final segments of the noun, in a relatively formal system, except for nouns denoting humans where a semantic system is used instead. But there are still traces of an implicative relationship. Thus, most nouns do not vary in number in French (there is addition of an *-s* in the orthography but it does not usually correspond to any actual phonological realization). But the variable class of nouns with a singular in *-al* and a plural in *-aux* only contains masculine nouns.

The implicative relationship between gender and inflectional classes is present for those languages where exponents marking number feature values on nouns can be shown to be related to the exponents marking gender and number on the targets for gender agreement, such as Arapesh languages of New Guinea or Niger-Congo languages. In these languages, one can show that the system of inflectional classes originated in the gender system.

8.2. Gender and inflectional class in Arapesh languages

Aronoff (1994:89-121) analyses gender and inflectional classes in two non-Austronesian languages from Papua New Guinea, Arapesh and Yimas. He shows in particular that gender and inflectional classes on nouns must be distinguished, although they are clearly linked. In particular, in Arapesh, inflectional classes on nouns are grouped in small groups whose members systematically correspond to one gender. Thus for gender IV, there are six inflectional classes on nouns, 4a to 4f. There is thus an implicational pattern between gender and inflectional class, and this pattern is unidirectional: every inflectional class on a noun unambiguously assigns that noun to a gender, but the reverse is not true [51] (Aronoff 1994:104-105). Therefore, there are more inflectional classes than genders. As for gender agreement, it is marked on proforms, adjectives, subject prefix agreement on the verb, and object agreement on the verb (see section 5.2.3. on object agreement in Arapesh). The forms for agreement can be shown to be formally similar, and similar to a number of markers on nouns. They thus probably evolved from the same forms, possibly through the grammaticalization of a classifier system or an article system. In addition, Arapesh presents a system of default gender agreement in the case of headless noun phrases, and for gender resolution in conjoined noun phrases of differing genders: in such cases, gender VIII agreement is used. Finally, inflectional classes themselves are mostly assigned following the phonological form of the singular of each noun, which determines to a large extent what the form of the plural will be.

Many languages of New Guinea present both gender and inflectional classes on nouns, which are very often traceable to similar forms originally. In this section, I analyse the system of gender and inflectional classes in Bukiyip, a close relative to Arapesh.

Bukiyip is an Arapesh language (Torricelli phylum) of New Guinea spoken in the East Sepik province of Papua New Guinea (Conrad & Wugigia 1991:1). Just like Arapesh originally, it has been described in terms of noun classes without making a distinction between gender and

51 There are some irregular nouns whose gender is not predicted by their inflectional class, but these are a small number.

inflectional classes; the number of genders devised for Bukiyip in the description available also implicitly only corresponds to the gender distinction made on the pronoun, not on other agreement targets, which complicates the picture (Conrad & Wugigia 1991). In this section, I separate the marking of inflectional class and gender in the language, establish the implicational patterns relating gender and inflectional classes, and speculate whether the gender system may have given rise to the inflectional class system historically. Although the system resembles that of Arapesh in some respects, there are some notable differences.

Table 126 summarizes the system of noun classes in Bukiyip, by providing the noun suffixes (i.e. the inflectional classes), one of the pronouns, the proximal pronoun, and the agreement forms on adjectives and verbs. In Table 126 the numeration of classes follows the 18 noun classes distinguished by the authors of the available description from which all the data are taken (Conrad & Wugigia 1991), although I show later that the number of classes is actually over 18.

	Noun suffix		PRO proximal		Adj suffix		Verb prefix	
	SG	PL	SG	PL	SG	PL	SG	PL
1a	-b	-bús	ébab	ébúsab	-bi	-búsi	b-	s-
1b	-n	-bús			-nali	-búsi	n-	s-
2	-bél	-lúb	éblab	éblalúb	-bili	-lúbi	bl-	bl-
3a	-g / -gú	-s / -as	égag	égsag	-gali / -gú	-gasi	g-	s-
3b	-ny	-gas			-nyi	-gasi	ny-	s-
3c	-gú	-gas			-gali	-gasi	g-	s-
3d	-m	-gas			-mi	-gasi	m-	s-
3e	-gúl	-gas			-nyi	-gasi	ny-	s-
4a	-k	-ou / -eb	okok	owour	-kwi	-wali	kw-	w-
4b	-k	-ial			-kwi	-wali	kw-	w-
4c	-k	-meb			-kwi	-wali	kw-	w-
4d	-k	-ib			-kwi	-wali	kw-	w-
4e	-mab	-meb			-bi	-bi	b-	b-
5a	-m / -bal	-s / -ipi / -bal	omom	éblab	-mi / -bali	-si / -ipi / -bali	m- / bl-	s- / p-/ bl-
5b	-m	-VpV(e/i)			-mi	-pi	m-	bl-
5c	-m	-bal			-mi	-bali	m-	bl-
6a	-n / -nú	-b	énan	ébab	-nali	-bi	n-	b-
6b	-nú	-b			-nali	-bi	n-	b-
7a	-n/ nú	-m	énan	ornom	-nali	-mi	n-	b-
7b	-nú	-m			-nali	-mi	n-	h-
7c	-nú	-lúh			-nali	-mi	n-	h-
8a	-ny / -l	-ch / -has	enyeny	echech	-nyi / -li	-chi	ny- / l-	ch-
8b	-V	-has			-nyi	-chi	ny-	ch-
8c	-l	-has			-nyi	-chi	ny-	ch-
8d	-m	-has			-nyi	-chi	ny-	ch-
9	-p	-s	énap	ésas	-pi	-si	p-	s-
10a	-l / -ny	-guh	élal	oguhogw	-li / -ny	-guhi	l- / ny-	hw-
10b	-ny ~ -l	-guh			-ny	-guhi	ny-	hw-
11a	-t / -tú	-gw	état	ogogw	-tali	-gwi	t-	gw
11b	-tú	-gw			-tali	-gwi	t-	gw
11c	-h	-g			-hi	-gali	h-	g-
12a	-hw	-lúh	ohohw	éhlah	-hwi	-lihi	hw-	hl-
12b	-hw	-lúh			-gali	-gasi	g-	s-
13a	-V_1h	-V_2h	éhah	ohoh	-hi	-h	h-	h-
13b	-p	-ih			-pi	-hi	p-	h-
14	-s	-s	ésas	ésas	-si	-si	s-	s-
15	-gún	-gún	égnag	ogohuh	-gúni	-gúni	gn-	gn
16	-has	-has	égúgún	égúgún	/	/	gn-	gn-
17	/	/	énan	ornom	-nali / -kwi	/	n- / kw-	/
18	/	-gún	éblab	éblalúb	/	-gúni	gn-	gn-

Table 126.Inflectional classes and genders in Bukiyip (data after Conrad & Wugigia 1991:10-14)

The system is striking in its complexity. Class 17 and 18 are recognized as separate on the ground that they concern personal names and place names respectively. They are characterized by defectiveness, and by the fact that personal names do not have specific endings as other classes do. They are thus the only classes that are not phonologically defined. In addition, class 18 takes the pronoun of class 2.

Table 126 shows a total of 38 inflectional classes on nouns (leaving out personal and place names), which in itself is a very complex system. They are characterized by patterns of alternation of the singular and the plural form of nouns. Some markers are ambiguous: thus *-b* marks the singular of class 1a but the plural of class 6. There are some patterns forming subclasses with default marking: thus four classes realize their singular as *-k* (class 4a to 4d), classes 8 and 16 mark their plural in *-has*, class 3b to 3e realize their plural as *-gas*. In Bukiyip, the presence of multiple defaults across classes for the same sets of feature values shows that the language infringes the No-Blur Principle.

Inflectional classes are distinguished from genders defined as agreement classes. The number of genders is problematic in that it depends on which target for gender agreement one looks at. For example, pronouns only distinguish 16 classes, because class 18 has the same set of pronouns as class 2, and class 17 has the same set of pronouns as class 7.[52] So one might be inclined to distinguish only 16 genders.

But if one looks at agreement patterns on adjectives and verbs, a very different picture emerges. There are nearly as many genders for these targets as there are inflectional classes. Taking into account both adjectival and verbal agreement, one can collapse together class 3b, 3e and 12b; class 4a, b, c, d; class 6a and 6b; 7b and 7c; 8b, c and d; and finally classes 11a and 11b. At the end of this operation one is still left with 30 agreement classes compared to the 38 inflectional classes and the 16 classes of pronouns. The maximal number of genders is thus 30 in Bukiyip.

The system is simplified by the fact that there are complex patterns of (partial) implication

52 This is true also of the pronominal forms which are not shown in the table.

between gender (and its expression in various targets) and inflectional classes. In particular, inflectional classes predict perfectly what gender they will belong to if one looks at pronominal gender. Similarly, the agreement classes on adjective and verb mostly predict what pronoun will be used.[53]

Inflectional classes seem to also predict rather well the agreement classes on adjectives and verbs. In addition, the form that agreement markers take on those targets can very often be related to the marker on the nouns. Table 127 presents the inflectional classes and agreement realizations of the first three major divisions of classes (the first three pronouns, and associated classes). Adjectival agreement in particular can be related very simply to most noun markers: in most cases it is the noun marker plus an additional final *-i*, in some cases an additional final *-ali*. Thus the noun plural marker *-gas* of class 3 corresponds to *-gasi* in adjectival agreement. For verbs there seems to be some default agreement in the plural with an omnipresent marker *s-*, but the forms of the singular retain the same form as the singular noun marker, only as a prefix, not as a suffix.

	Noun suffix		Adj suffix		Verb prefix	
	SG	PL	SG	PL	SG	PL
1a	-b	-bús	-bi	-búsi	b-	s-
1b	-n	-bús	-nali	-búsi	n-	s-
2	-bél	-lúb	-bili	-lúbi	bl-	bl-
3a	-g / -gú	-s / -as	-gali / -gú	-gasi	g-	s-
3b	-ny	-gas	-nyi	-gasi	ny-	s-
3c	-gú	-gas	-gali	-gasi	g-	s-
3d	-m	-gas	-mi	-gasi	m-	s-
3e	-gúl	-gas	-nyi	-gasi	ny-	s-

Table 127. The first three major classes in Bukiyip (Conrad & Wugigia 1991:10-14)

In any case, there is a relatively strict relationship in form between noun inflectional markers and agreement markers on adjectives and verbs. One can thus make a hypothesis as to the diachronic origin of the system. Because of the formal link between inflectional class and agreement class, one can posit that the two systems originated in the same markers, and subsequently became differentiated due to sound change and analogy. It is also possible to compare

53 There are some exceptions such as class 12a, whose agreement pattern is identical to class 3b, but takes a different pronoun. The implicational pattern is thus not a hundred per cent predictive.

the classes in Bukiyip with the (more limited number of) classes of Arapesh to reconstruct an earlier system. A major sound correspondence is /l/ in Bukiyip corresponding to /r/ in Arapesh. Most Arapesh classes have clear reflexes in Bikiyip. I propose that these common classes can be understood as being the original classes before more allomorphy developed. Table 128 shows the corresponding inflectional classes of Arapesh and Bukiyip.

BUKIYIP			ARAPESH
1a	-b / -bús	-by̥ / -bys	1
2	-bél / -lúb	-bør / -ryb	2
3a	-g~-gú / -s~-as	-ag / -as	3a
3c	-gú / -gas	-g / -gas	3b
4a	-k / -ou	-ku / -u	4f
4c	-k / -meb	-ku / -meb	4a
4d	-k / -ib	-ku / -ib	4c
5b	-m / VpV	-m / ipi̥	5
6a/b	-n~-nú / -b	-n / -b	6
7a/b	-n~-nú / -m	-n / -m	7
8a	-ny / -ch	-iñ / -ish	8a
8b	-V / -has	-V / -Vhas	8b
9	-p / -s	-pu / -s	9b
10a/b	-ny~-l / -guh	-r / -guh	10
11a/b	-t~-tú / -gw	-t / -gu̥	11b
12a/b	-hw / -lúh	-uh / -ruh	12
13a	-V_1h / -V_2h	-ah / -eh	13

Table 128. Comparison of Bukiyip and Arapesh inflectional classes (Conrad & Wugigia 1991:10-12; Aronoff 1994:91)

The link between classes in Arapesh and in Bukiyip is even clearer when one compares the agreement forms, as shown in Table 129 for adjectival agreement. These probably show the original set of agreement markers which were also used as noun class markers on nouns.

BUKIYIP			ARAPESH
1a	-bi / -búsi	-byi / -bysi	I
2	-bili / -lúbi	-børi / -røbi	II
3a	-gali / -gú	-gi / -gasi	III
4a-d	-kwi / -wali	-kwi / -ui	IV
5b	-mi / -pi	-mi / ipi	V
6	-nali / bi	-ni / -bi	VI
7	-nali / mi	-ni / -mi	VII
8	-nyi / chi	-ñi / -shi	VIII
9	-pi / -si	-pi / -si	IX
10a	-li~-nyi / -guhi	-ri / -ruhi	X
11a/b	-tali / -gwi	-ti / -gwi	XI
12a	-hwi / -lihi	-whi / -guhi	XII
13a	-hi / -h	-hi / hi	XIII

Table 129. Bukiyip and Arapesh adjective agreement (Conrad & Wugigia 1991:10-12; Aronoff 1994:92)

One can thus reconstruct a uniform system, where the markers for gender on agreement targets and the markers of inflectional class originate in the same type of marker, be it an article or a classifier, which was apposed to nouns, verbs and adjectives, and progressively grammaticalized. The inflectional class system thus seems to have exactly the same origin as the gender system, as if it had first been an overt gender system, which underwent subsequent reanalysis. Thus for example, one can trace the fact that class 12b and class 3a/b have the same agreement on adjectives and verbs to a subsequent change, some nouns of class 12 acquiring a different gender agreement. Similarly, there seems to have been some analogical extension of the marker *-ny* for singulars from originally class 8 (the default class in Arapesh) to class 10 and 3 in Bukiyip. It is not so surprising that forms belonging to the default class should get extended to other classes.

8.3. Miraña

The very specific system of gender exhibited by Miraña can shed some light on how a gender system can give rise to a system of inflectional classes. Miraña (also called Bora) is a Boran language from South America spoken in the Amazon region of Brazil (Seifart 2005; Hammarström et al. 2017). The language has a very peculiar gender/classifier system that can help us disentangle

how inflectional classes and gender systems can arise at the same time, or rather, from the same source. Miraña shows noun classes on nouns, and gender agreement with similar forms on a number of targets inside the noun phrase. The markers used for noun classes and for gender agreement are most of the time identical. But it does not present inflectional classes, unless one considers the exponent of the noun classes marked on nouns as thematic elements: the marking of plurality, for example, does not vary from noun to noun. It still overtly marks the gender class of a noun, in addition to marking it through agreement. Miraña shows a complex gender system which arose from the grammaticalization of a classifier system.

Miraña presents a system of 72 noun classes. These classes are distinguished through noun class markers suffixed to the nouns. There are two sets of markers, six general class markers, which encode distinctions such as animacy, sex and number, and can replace specific markers in agreement. In addition there are 66 specific markers which mostly encode shape distinctions and are used for inanimates. As Seifart summarizes the system, "general class markers are a small set of forms that encode animacy, sex, and number. Specific class markers are a large and internally diverse set consisting of monosyllabic and polysyllabic forms. Within specific class markers, a core set of monosyllabic forms with relatively broad meanings can be identified. These are the forms that are most frequently used in nouns and as agreement markers." (Seifart 2005:103). Because there is agreement in noun classes, the markers on the nouns can be shown to be overt gender markers.

Agreement in noun class (gender) in Miraña is shown on a number of targets, including main clause predicates and nominal expressions such as pronouns, numerals and relative clauses, which agree in noun class and mark agreement on their predicate (Seifart 2005:157). Miraña thus presents a system of gender, shown through agreement, in addition to the overt marking of noun class on the noun. Most of the markers used in agreement are identical or similar to those used to mark noun class on nouns, with a few exceptions. Example 3 shows such agreement inside the noun phrase, with a numeral, Example 4 shows agreement with similar forms with a predicate. Example 5 shows that in some cases the forms used in agreement are not identical to the form used to mark noun

class.

3) tsa-:pi gwa-hpi
one-CL.MASC.SING human-CL.MASC.SG
'one man' (Seifart 2005:158)

4) kátɯ́:βɛ-i dɯrɯ́-i
fall-CL.1D.MEDIUM wax-CL.1D.MEDIUM
'It (stick-shaped) fell, the candle' (Seifart 2005:158)

5) kátɯ́:βɛ-:bɛ gwa-hpi
fall-CL.MASC.SG human-CL.MASC.SG
'He fell, the man' (Seifart 2005:158)

But Miraña really shows a system of mostly alliterative concord (although there are cases where the markers are different, as shown in Example 5 above), with the same markers repeated on both controller nouns and targets for agreement. This can be shown for nouns which can take more than one gender. For example, the noun for avocado can be derived with two different class markers. When a given form is used, agreement can only be with that same form, and cannot be expressed with the other possible class marker, as shown in Examples 6 and 7:

6) tsa-ʔba kó:hɯ-ba
one-CL avocado-CL
'one avocado (fruit)'

7) tsa-ʔo kó:hɯ-ʔo
one-CL avocado-CL
'one avocado (fruit)' (Seifart 2005:160)

Alternatively, inanimate noun classes can see their agreement marker replaced by a general inanimate gender marker *-nɛ*. It is possible to combine this marker on some targets while others bear the expected class marker (Seifart 2005:167-168). But if the noun itself is marked with the general inanimate class marker, all targets have to agree with this same marker.

The Miraña system clearly arose from a previous system of classifiers encoding shape of objects and animacy distinctions. This system of classifiers has grammaticalized and expanded to give rise to a gender system characterized by agreement in noun class with a number of targets. It has not developed as a system of inflectional classes in nouns, except if one considers the class markers as thematic elements. Miraña thus shows a pathway for the acquisition of both gender and inflectional classes from the same source, which can be paralleled in other languages such as Arapesh languages or Niger-Congo languages: the marking of the same gender based distinction on both nouns and agreement targets. This double marking can under some circumstances give rise to inflectional classes: the marking of class on nouns has to be linked with inflection (for example through the marking of number on nouns, which is the case in Miraña), and further evolution has to fuse the noun class marker and the number marker for it to give rise to inflectional classes. Even in the case of Miraña, one could consider the class markers as thematic markers dividing nouns into as many inflectional classes, even though the plural marker does not vary.

8.4. Gender and inflectional class in Niger-Congo languages

Niger-Congo languages exemplify inflectional class creation from a gender system. It can be shown that the two systems share a common origin in those languages because they share most of their markers in some languages, and the proto language is reconstructed with such shared markers. But inflectional classes and gender agreement evolve as two separate systems in subsequent stages of the development of some language families within Niger-Congo, particularly the Gur languages.

8.4.1 The gender system of proto-Niger-Congo

Most major groups of Niger-Congo languages present systems of nominal classification (Good 2012; Kießling 2013), with the exception of Mande languages (although see Vydrin 2006 who analyses possible traces of an earlier classification system in Mande languages). It is thus generally accepted that proto-Niger-Congo also had a system of gender and noun classes comparable to what

is seen nowadays in Bantu languages: nouns mark inflectional classes formed from a marker for singular and a marker for plural, and the same markers are used in agreement to mark a number of targets inside and outside the noun phrase. Noun classes and genders have been reconstructed for most primary groups of Niger-Congo, and the forms of these reconstructions show striking similarity between primary groups. Kießling (2013) synthesizes those reconstructions, but no attempts have been made until now to reconstruct a gender system further for proto-Niger-Congo. Beyond the forms used for each gender, there is also similarity in the semantic core of each gender: one can reconstruct a number of typical nouns which belong to each of the genders. Thus gender 1/2 is very firmly established in all groups with a semantic core relating to humans, and with similar forms in the different groups: **gu* for the singular in Kordofanian; **u-gu-* for the singular and *a-ba-* for the plural in Atlantic; **-ʊ* or **-a* in the singular and **-ba* in the plural for Gur; singular **o-*, plural **ba-* in Kwa; singular **ù-*, plural **bà-* for Benue-Congo; and finally singular **mò-* plural **bà-* for Bantu (Kießling 2013:46). Similarly, gender 3/4 typically consists of nouns meaning tree, tree names, rope, broom, road, and fire, while gender 6a concerns liquids (Kießling 2013:46).

Kießling (2013) shows convincingly how such a system could have arisen in the first place in proto-Niger-Congo. His reconstruction is interesting for the purpose of this chapter because he posits an identical origin for the markers on nouns forming inflectional classes, and for the markers of agreement. His reconstruction is based on the fact that grammaticalization can operate in cycles, and that a number of Niger-Congo languages have developed, in addition to their gender system, a system of (numeral) classifiers or a very developed system of class-terms (of the type of English X-berry). This, together with the fact that Niger-Congo genders revolve around a common semantic core that can be said to originate in a semantics-based classifying system leads him to posit that the system of noun classes originated in a system of classifiers. Thus, class terms and measure terms in quantifying expressions pave the way for the creation of a classifier system which can be extended to the whole lexicon. The terms used for classifiers are nouns: body part terms, measure terms, etc similar to those attested in contemporary numeral classifier systems in Africa. Once the system of

classifiers has expanded to cover all nouns, it can then be extended syntactically to other domains, which creates agreement: NP-internally this is mostly done through the used of classified demonstratives, outside the NP, through the use of anaphorical expressions which are also classified. Marking on nouns can either be an extension of such a system of anaphoric marking through the use of classified articles appearing adjacent to the noun (a scenario favoured by Greenberg (1978), see the introduction to this chapter), or alternatively it can originate directly in the grammaticalization and formal fusion of a classifier appearing adjacent to the noun. The distinction in forms for singular and plural could be explained by the fact that a number of basic and generic terms in African languages frequently show suppletion for number. A simple example is the opposition person / people which could have given rise to the two unrelated markers of class 1/2 for singular and plural. In any case, this reconstruction means that inflectional classes and genders originate from the same system of classifiers (or alternatively following Greenberg the system of inflectional classes is a direct product of the gender system), through a process of grammaticalization.

In a number of Niger-Congo languages (and this can also be posited for proto-Niger-Congo), there is a very robust pattern of implication between the morphology of inflectional classes on nouns and the agreement classes. The typical Bantu language for example is deemed to present exactly the same number of genders as the number of inflectional classes on nouns, and each gender corresponds very precisely to an inflectional class. In fact, it is very hard to find a language where such alliterative agreement is consistently made for all values and all targets (Greville Corbett p. c.). In such languages, the pattern of implication is thus bidirectional, which seems to be fairly rare for implicative morphology: genders predict inflectional classes, but inflectional classes also predict genders. This bidirectionality may be an indication that the two systems have a common origin. In addition, the markers used for inflection on nouns and for agreement on the target tend to be formally identical. Thus in Example 8 and 9 from Swahili, the markers of gender 7/8 are *ki-* in the singular and corresponding *vi-* in the plural for both controller nouns and the targets for agreement.

Such systems can be said to be maximally overt gender systems.

8) Ki-kombe ki-dogo ki-moja ki-me-vunjika
   7-cup 7-small 7-one 7-PERF-be.broken
   'One small cup is broken' (Kießling 2013:45)

9) Vi-kombe vi-dogo vi-wili vi-me-vunjika
   8-cup 8-small 8-one 8-PERF-be.broken
   'Two small cups are broken' (Kießling 2013:45)

In this section, I show that although the two systems, inflectional classes on nouns and genders, in some languages form a system of mutual implications, the two systems are still completely independent, although they use the same markers. Thus the markers do not involve repetition of identical lexical entries, but the markers on nouns and the markers on agreement targets realize different features. In fact, one system can be lost without the other being affected by this loss. The Gur languages offer a good example of the loss of agreement, and by consequence of the gender system of the language, without the system of inflectional classes on nouns being lost as well.

8.4.2 Losing gender, keeping inflectional classes: the Gur languages

The Gur group of Niger-Congo languages presents a peculiarity with respect to its relationship between gender and inflectional classes. Noun classes marked on nouns, and a system of agreement on pronouns and some other parts of speech can be reconstructed in proto-Gur (Miehe et al. 2007). Just like proto-Niger-Congo, the protolanguage probably had a system of alliterative concord, with overt marking of gender on nouns forming inflectional classes based on the number feature. But this state of affairs has changed dramatically over time, and most Gur languages nowadays present a system where noun classes are preserved on nouns, but where agreement has nearly disappeared or has been remodelled on a system opposing animates to inanimates. In this section, I first describe a language that has kept some agreement classes with pronouns corresponding to the classes distinguished on nouns, Cerma; then I describe the situation in Dagaare which has nearly lost all

concord, and Chakali, which had completely lost its gender system before reinnovating a new one, not based on noun classes. What the evidence from the Gur languages shows is that, although the marking of inflectional classes on nouns and the agreement marking of gender on other parts of speech have the same origins, and although they are initially linked through very strong implicative patterns (one to one correspondence between inflectional classes on nouns and genders shown through agreement), one system can be remodelled or lost without affecting the other system. The system of noun class marking on nouns is thus not simply a system of overt gender marking on nouns: in such cases, it is rather a fully independent system of inflectional classes, with a common origin with the gender system, and with some degree of implicative relations linking the two systems, which can evolve completely independently. In the Gur languages, the system of gender is lost or remodelled, but the system of inflectional classes on nouns remains very stable.

8.4.2.1. Cerma

Cerma is a Gur language spoken in Burkina-Faso and Ivory Coast (Miehe 2007). It is remarkable in the Gur languages in that it has kept a system of agreement on pronouns corresponding exactly to the noun classes marked on nouns. Concord is marked on the full range of pronouns of the language, including subject pronouns, which also function as possessive markers, object pronouns, demonstrative pronouns, indefinite pronoun and interrogative pronouns, but it is not marked on other targets in the noun phrase (Miehe 2007:60-61). Table 130 indicates the affix markers and the corresponding subject pronouns (in bold), as well as the corresponding proto-Gur noun classes numbered according to the reconstruction in Miehe & al. (2007). The lines in the chart mark the link between singular and plural forms forming classes and gender, the dotted line indicates a subclass with very few members. Gender and noun classes are still linked by a strong implicative pattern: a given gender corresponds exactly to one inflectional class.

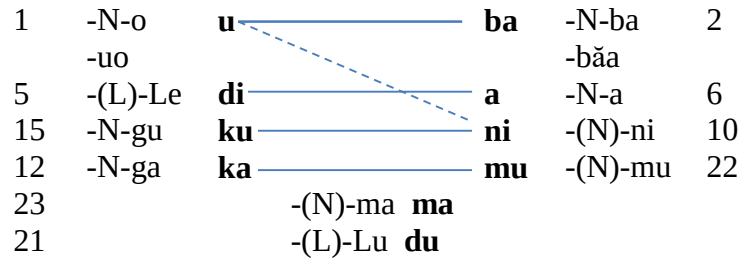


Table 130. The noun class system of Cerma (Miehe 2007:24)

In Cerma, the markers on nouns are nearly systematically prenasalized, except in one of the variants of class 1/2. The language marks seven inflectional classes on nouns (two defective), and seven genders.

Cerma is a conservative language, where the state of affairs present in proto-Niger-Congo is preserved: noun classes are marked on nouns, forming inflectional classes varying according to the feature number; there is gender agreement on pronouns exclusively (a restriction compared to the proto-language). Agreement classes correspond strictly to inflectional classes on nouns in a bidirectional implicative relationship. And finally the markers used on nouns and on pronouns show clearly a similar form, which confirms their common origins.

8.4.2.2. Dagaare

Dagaare is part of the Dagara dialect cluster spoken near the Volta River in Burkina-Faso and Ghana (Miehe 2012; Bodomo 1997). The language shows a more extended number of inflectional classes on nouns (Table 131), but a dramatically reduced system of agreement classes. In addition, it also shows lexically defined inflectional classes on adjectives, rather than agreement. The lines in the chart indicate the pairings of singular and plural forms which form inflectional classes.

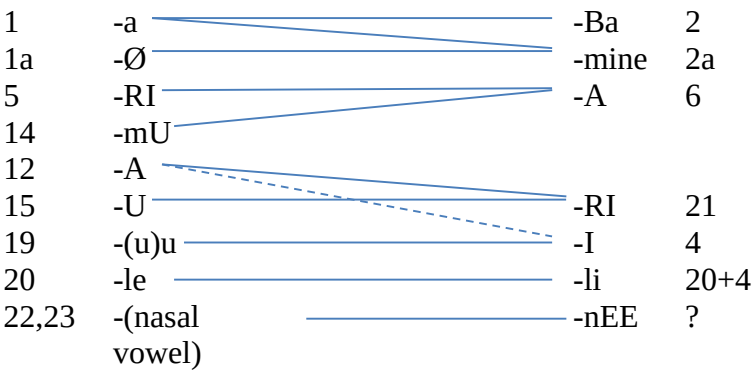


Table 131. Inflectional classes in Dagaare (Miehe 2012:251)

Dagaare presents ten inflectional classes marked on nouns (Bodomo 1997), which can be reconstructed to noun classes in proto-Gur which took part in the gender system. These markers were initially the markers of an overt gender system. Because agreement classes were reduced in Dagaare like in a number of Gur languages, what is left is inflectional classes.

The agreement system of Dagaare is extremely reduced, and does not correspond any longer to the system of noun classes. Because most animates are in gender 1/2, the markers for the animate gender are those of this class, while new markers have been innovated for inanimates, which no longer distinguish noun class. These markers do not distinguish between singular and plural in Dagaare. Some dialects of the Dagara cluster have lost all gender distinction, such as in Wandara and Birifor. Table 132 shows the pronouns for animate and inanimate for the different dialects:

	SG.ANIM	PL.ANIM	SG.INAN	PL.INAN
Dagaare	ʊ	ba	a	a
Dagara-Lobr	ʊ	bɛ	a	a
Dagara-Wile	ʊ	ba	ʊ	a
Wandara	ʊ	ba	ʊ	ba
Birifor	ʊ	ba	ʊ	ba

Table 132. Pronominal agreement in Dagara dialects (Miehe 2012:253)

In Dagaare, adjectives also present inflectional classes that seem derived from the same classes as those marking nouns. They are lexically specified, as shown in Example 10, and do not depend on the gender of the noun (Miehe 2012:265). The presence of class markers on adjectives probably indicates that in the proto-language there was agreement in gender with adjectives, but this agreement was lost and gender markers lexicalized with adjectives, probably based of their frequency of cooccurrence. They are mostly found with gender 12/21, but some are also in gender 15/21.

10) 12/21 zɪɛ / zɪʊrɪ

red.SG / red.PL

'red' (Miehe 2012)

Under some analyses, adjectives in fact form compounds with the noun, and plural is only

marked after the string of noun and adjectives, not on the noun itself. But in any case, there is no agreement in gender between the noun and the adjective, as shown in Example 11 and 12:

11) yi-zeɛ

house-red

'red house' (Bodomo 1997:49)

12) yi-zee-re

house-red-PL

'red houses' (Bodomo 1997:49)

8.4.2.3. Chakali

In Chakali (Brindle 2009), the agreement classes have completely disappeared, leaving only inflectional classes which mark pairs of singular and plural forms of the nouns by the means of suffixes. There are more classes than suffixes, which means that the distribution of affixes is what defines classes. There are five major classes, and a number of smaller, less productive classes (Brindle 2009:86-87). According to Brindle (2009:87), the forms taken by the affixes can be related to reconstructed forms in proto-Gur as well as forms taken by gender classes in related languages. In addition to the classes cited, there are also a few irregular nouns, and Brindle indicates that there is some degree of dialectal variation in the assignment of nouns to classes, which is a common situation for inflectional classes, but less so for gender systems based on semantics. Chakali has recently innovated a new gender system that does not seem to have any link with the previous system, and marks two genders based on animacy distinctions on third person plural only, which is typologically unusual (Brindle 2009).

8.4.3 Stable inflectional class in Niger-Congo

Inflectional classes seem to be more robust than gender marked through agreement, at least in some branches of Niger-Congo languages. Thus in Gur languages, agreement is highly reduced or

often lost all together in favour of a new system of gender marking based on animacy. But the marking of classes of singular / plural pairings on nouns is still present and does not seem to be lost in any language of the family. The Kwa branch of Niger-Congo seems to have lost both gender and inflectional classes, but there are often some remnants of marking on nouns (Good 2012). Sometimes, there can be changes in the noun class system which do not affect the marking of gender. Thus Good (2012) mentions the cases of Noni and Mundabli, where the marking of inflectional classes has evolved towards non-segmental marking on nouns. Distinctions of class are still present in the system though, even if it is not through prefixal marking. On the contrary, there does not seem to be any language where inflectional classes are lost while the system of concord is preserved. The system of inflectional classes thus seems to be more robust than the system of gender marking. The two systems have the same origins, the system of inflectional classes originating in the marking of gender overtly on nouns, but the two systems are clearly independent.

8.5. Conclusion

This chapter has shown that there is often an implicative relationship between gender and inflectional classes. This relationship is clearly present in some European languages, including Spanish and Russian, although the two systems do not originate in the same original system of gender marking. At the opposite end, Niger-Congo languages and Arapesh languages show clear instances where the system of inflectional classes and the system of gender originate in the same set of markers. In both of the latter cases, the inflectional class is defined on the feature number, which seems to be particularly prone to fusing with gender and to create inflectional classes from a gender system. It thus seems that systems of inflectional classes can originate in systems of overt gender-marking on nouns.

Similarly to the preceding chapter on alienability distinctions, when a system of inflectional classes evolves from a system of gender marking, this is due to various processes of

grammaticalization, but such processes apply to an already fully formed system of classes. The overt marking of gender on nouns, when applied to nouns which also mark a range of inflectional features, develops into a system of inflectional classes. This system is peculiar in that the members of each inflectional class are also the members of a given gender. The form of the classes, and their membership, is thus fully determined by the form and membership of the gender classes. The two systems end up being related to each other by patterns of implication. Such patterns can evolve in diachrony, and the evolution of both systems in the Gur languages shows that, although originating in the same distinctions, the two systems are fully independent.

Following the main argument in this thesis, that inflectional classes arise from the cooccurrence of the marking of inflectional values with a distinction made between groups of lexemes having some other peculiarities, here, this peculiarity is gender, which already partitions the lexicon of nouns into various classes, which are inherited by inflectional classes when they arise from a system of gender.

9. Conclusion

This thesis has set out to understand how inflectional classes arise in natural languages, what processes are involved, and what conditions need to be met for such autonomously morphological objects to be created. In order to do so, it has been necessary to first define what inflectional classes are, and to clarify existing definitions. This has confirmed that inflectional classes have to be understood as classes of lexemes, not as classes of stems. These classes are characterized by a specific range of allomorphs for realizing feature value pairs. Inflectional classes share a content paradigm, but differ in their form paradigm, albeit minimally. Separate classes have to be recognized when at least one form differs, although such a case is highly noncanonical, as canonical classes should differ systematically in the forms realizing feature values. Similarly, there is no theoretical difference between irregulars and classes with more members, except for the fact that in the case of irregulars the set of realizations does not serve as a model for inflecting other lexemes. Irregulars are thus only classes with one member. Inflectional classes normally arise within a given word class, but under some conditions they can be shared across word classes. Examples of such sharing given in the literature have been shown to be erroneous, but other examples can be found, in particular adjectives and possessive pronouns in Latin. One consequence of defining inflectional classes as lexical classes is that heteroclites have to be recognized as a separate inflectional class, but one which inherits its realization from two other classes instead of one in an inheritance hierarchy model.

Inflectional classes emerge as having very diverse origins. I have confirmed here that the main processes involved in the creation of inflectional classes are sound change, grammaticalization and reanalysis. These three processes had been shown to be implicated in the creation of inflectional classes in Germanic languages (Dammel 2011). This thesis extends the range of languages where such developments are attested, adding Romance languages, Basque, Austronesian languages, and non Austronesian languages of New Guinea. Furthermore, I have clarified and expanded the range of possible origins for inflectional classes in the case of grammaticalization. Until now, it was

understood that grammaticalizations of the agglutination type could give rise to inflectional classes. I have added the grammaticalization of auxiliary systems as a possible source, as well as successive cycles of grammaticalization of pronouns. This last development in particular has enabled me to present a rare type of inflectional class, where classes are distinguished by the linearization of affixes: although the forms are similar, in Arapesh, one class is marked by suffixes while the other is marked by prefixes.

Beyond confirming these scenarios, I also show that different types of origins for inflectional class systems will give rise to systems which differ systematically in their form. Thus, sound change will generally split an existing class into two classes which partially inherit forms from the previous one, while reanalysis will generally create more classes at once.

Other types of inflectional class creation which have not been recognized in the literature have also been analysed. This is in particular the case for morphological phenomena considered as noncanonical inflection, which have to be considered as inflectional classes because of the definition of classes as lexical classes sharing a content paradigm but not a form paradigm. Two such phenomena are heteroclisis and deponency.

Finally, inflectional classes can also develop from existing systems presenting lexical classes, mainly systems of alienability distinctions in possessive paradigms, and systems of gender when they are marked overtly on nouns. In such cases, the membership of classes and the number of classes, as well as the inflectional material realizing the set of feature values are inherited from a previous system: Niger-Congo languages mark the same number of inflectional classes as they mark genders, initially. But the two systems can be shown to be fully independent, as they may evolve in opposite directions. Gur languages lose gender agreement while retaining inflectional classes on nouns; a number of new inflectional classes develop in Bukiyip by analogy or sound change after they have been initially created from a gender system.

There is no overarching principle which can be used to find sources for inflectional classes. Any distinction between sets of lexemes added to inflectional marking can give rise to a system of

inflectional classes. This is what explains why such a varied number of possible origins can be found. I have not examined in any detail phenomena such as analogy, which does not seem to play a role in the creation of new classes, or at least I could find no examples of it. There is thus some limitation as to which phenomena can produce inflectional classes in diachrony. Analogy usually tends to bring a given paradigm closer to an existing one: if two paradigms diverge in a subset of forms, analogy will have the tendency to bring one closer to the other, and may ultimately make the two paradigms identical. It is thus more a force that works towards the convergence of classes, rather than their subdivision, and thus it is not a priori a possible source for inflectional classes.

Some questions have been left for future research. I have shown that different sources give systems of inflectional classes with very different forms. Some changes only split existing classes or inflection proper: this is typically the case with the creation of classes following sound change. Other changes apply to a larger number of classes. Alienability distinctions give rise to a very peculiar system of classes where there is generally a large open class, and a number of small membership, closed classes, sometimes also with heteroclite classes. It remains to be seen to what extent one could predict the origins of a given system following its form, because in actual systems, subsequent changes, either sound change or analogy, often obscure initial developments.

An indication as to the origins of some inflectional class systems may be found in the extramorphological conditioning of classes which remains after their evolution. Actual classes are noncanonical on the lines of one of the extramorphological conditionings examined by Corbett (2009): canonical classes should not be determined by phonological, semantic or syntactic factors. In fact, classes will often show only one such conditioning, but it will be a partial conditioning: there is some implicational link for example between phonology and inflectional classes in Biak, but the phonology does not determine the membership of classes perfectly. This conditioning is an indication that classes in Biak originally evolved from a phonologically conditioned alternation, probably through sound change. Similarly, classes originating in a gender system may show some

level of semantic conditioning for those gender classes with semantic assignment. It is thus possible, in the presence of a synchronic system, to start investigating its possible origins following its noncanonical form.

I have not asked which features favour the development of inflectional classes, or whether any given feature may favour such developments. Similarly, inflectional classes seem to emerge inside a given word class. But the question of which word classes favour the emergence of inflectional classes has not been asked. It is highly possible that they develop mostly in large membership, open word classes such as verbs and nouns, as most of the examples presented in this thesis come from these two word classes, but these are also the word classes present in all languages, and thus statistically more prone to presenting inflectional class development. Not all languages have an adjective class, but inflectional classes are also attested in adjectives, at least for languages where they form an open class, such as Latin or the Romance languages more generally.

Finally, I have left for further research the link between various autonomously morphological phenomena. One could imagine that the presence in a given language of morphomes, for example in the form of root allomorphy distributions, could favour the emergence of the other main morphomic object, namely inflectional classes. It is very often difficult to know whether such phenomena appeared before or after inflectional classes. Romance languages and Latin for instance, present both phenomena. Work done on the Romance languages (Maiden 2018) shows some patterns of implication betweel inflectional classes and morphomic distributions of root allomorphy. For instance, there is a general tendency for first class verbs in Romance to resist patterns of root allomorphy. Other types of patterns of root allomorphy, such as the L-pattern in Ibero-Romance, are stricly associated with non-first conjugation class. Future research extended to the Indo-European family may shed some light on the particular question of the diachronic relationship of morphomic objects.

References

Ackerman, Farrell & Robert Malouf. 2013. Morphological organization: the low conditional entropy conjecture. *Language* 89 (3), 429-464.

Ackerman, Farrell & Robert Malouf. 2015. The No Blur Principle effects as an emergent property of language systems. *Proceedings of the 41st Annual Meeting of the Berkeley Linguistic Society*, 1-14.

Ackerman, Farrell & Gregory Stump. 2004. Paradigms and periphrastic expression: A study in realization-based lexicalism. In Louisa Sadler & Andrew Spencer (eds.), *Projecting Morphology*. Stanford: CSLI, 111-157.

Anderson, Stephen R. 2015. The morpheme: its nature and use. In Matthew Baerman (ed.), *The Oxford handbook of inflection*. Oxford: Oxford University Press, 11–33.

Ariztimuño, Borja. 2013. Finite Verbal Morphology. In Mikel Martínez-Areta (ed.) *Basque and Proto-Basque. Language-Internal and Typological Approaches to Linguistic Reconstruction*. Frankfurt am Main: Peter Lang.

Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge, Mass: MIT Press.

Baayen, R. Harald, McQueen, J., Dijkstra, T., and Schreuder, R. 2003. Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In Harald Baayen & Robert Schreuder (eds), *Morphological structure in language processing*, Berlin: De Gruyter Mouton, 355-390.

Bach, Xavier. 2017. Alienable possession in Biakic. Ms, University of Oxford.

Bach, Xavier & Louise Esher. 2015. Morphological evidence for the paradigmatic status of infinitives in French and Occitan. In Dag T.T. Haug (ed.), *Historical Linguistics 2013. Selected papers from the 21st International Conference on Historical Linguistics, Oslo, 5-9 August 2013*. Amsterdam: John Benjamins, 135–154.

Baerman, Matthew. 2007. Morphological Typology of Deponency. In Matthew Baerman, Greville Corbett, Dunstan Brown & Andrew Hippisley (eds.), *Deponency and Morphological*

Mismatches. Oxford: Oxford University Press, 1-19.

Baerman, Matthew. 2012. Paradigmatic chaos in Nuer. *Language* 88 (3), 467-494.

Baerman, Matthew (ed.). 2015. *The Oxford handbook of inflection*. Oxford: Oxford University Press.

Baerman, Matthew, Dunstan Brown & Greville Corbett. 2005. *The Syntax-Morphology Interface: a Study of Syncretism*. Cambridge: Cambridge University Press.

Bauer, Laurie. 2004. The function of word-formation and the inflection-derivation distinction. In Henk Aertsen, Mike Hannay & Rod Lyall (eds.), *Words in their Places. A Festschrift for J. Lachlan Mackenzie*. Amsterdam: Vrije Universiteit, 283–292.

Berg, René van den. 2009. Possession in South Halmahera-West New Guinea: typology and reconstruction. In Alexander Adelaar & Andrew Pawley (eds.), *Austronesian historical linguistics and culture history: a festschrift for Robert Blust*. Canberra: Pacific Linguistics, 327–357.

Berg, René van den & Takashi Matsumura. 2008. Possession in Irarutu. *Oceanic Linguistics* 47 (1), 213-222.

Bjerrome, Gunnar. 1957. *Le Patois de Bagnes (Valais)*. Stockholm: Almqvist & Wiksell.

Blevins, James P. 2005. Word-based declensions in Estonian. *Yearbook of Morphology 2005*, 1-25.

Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 42 (3), 531–573.

Blevins, James P. 2015. Inflectional Paradigms. In Matthew Baerman (ed.), *The Oxford Handbook of Inflection*. Oxford: Oxford University Press.

Blevins, James P. 2016. *Word and Paradigm Morphology*. Oxford: Oxford University Press.

Blevins, Juliette. 1994. A phonological and morphological reanalysis of the Maori passive. *Te Reo* 37, 29-53.

Blevins, Juliette. 2008. Consonant epenthesis: Natural and unnatural histories. In Jeff Good (ed.), *Linguistic Universals and Language Change*. Oxford: Oxford University Press.

Blevins, Juliette & Andrew Garrett. 1998. The origins of consonant-vowel metathesis.

Language 74, 508-56.

Bodomo, Adams. 1997. *The Structure of Dagaare*. Stanford: CSLI.

Bonami, Olivier. 2015. Periphrasis as collocation. *Morphology* 25 (1), 63-110.

Bonami, Olivier & Gilles Boyé. 2003. Supplétion et classes flexionnelles dans la conjugaison du français. *Langages* 152, 102-126.

Bonami, Olivier & Gilles Boyé. 2007. Remarques sur les bases de la conjugaison. In E. Delais-Roussarie & L. Labrune (eds.), *Des sons et des sens. Données et modèles en phonologie et en morphologie*. Paris: Hermès-Lavoisier, 77-90.

Bonami, Olivier & Pollet Samvelian. 2009. Inflectional periphrasis in Persian. In Stefan Müller (ed.), *Proceedings of the HPSG 2009 conference*. Stanford: CSLI, 26-46.

Booij, Geert. 1994. Against Split Morphology. In Geert Booij & Jaap Van Maarle (eds.). *Yearbook of Morphology 1993*, 27–49.

Booij, Geert. 1996. Inherent versus contextual inflection and the split morphology hypothesis. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1995*, 1–16.

Brindle, Jonathan A. 2009. On the identification of noun class and gender systems in Chakali. In Masangu Matondo, Fiona McLaughlin & Eric Potsdam (eds.). *Selected proceedings of the 38th Annual Conference on African Linguistics*. Somerville: Cascadilla Proceedings Project, 84-94.

Brown, Dunstan, Marina Chumakina, Greville Corbett, Gergana Popova & Andrew Spencer. 2012. Defining 'periphrasis': key notions. *Morphology* 22 (2), 233-275.

Brown, Dunstan & Andrew Hippisley. 2012. *Network morphology: a defaults-based theory of word structure*. Cambridge: Cambridge University Press.

Bugenhagen, Robert. 1995. *A grammar of Mangap-Mbula: an Austronesian language of Papua New Guinea.* (Pacific Linguistics. Series C, 101). Canberra: Australian National University.

Bybee, Joan L. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar. Tense,*

aspect, and modality in the languages of the world. Chicago: The University of Chicago Press.

Calabrese, Andrea. 2011. Metaphony in Romance. In Mark van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*. Oxford: Blackwell, 2631-2661.

Cameron-Faulkner, Thea & Andrew Carstairs-McCarthy. 2000. Stem alternants as morphological signata: Evidence from blur avoidance in Polish nouns. *Natural Language & Linguistic Theory* 18 (4), 81-835.

Cappellaro, Chiara. 2013. Overabundance in diachrony: a case study. In Silvio Cruschina, Martin Maiden & John Charles Smith (eds.), *The boundaries of pure morphology: Diachronic and synchronic perspectives*. Oxford: Oxford University Press. 209-220.

Carstairs-McCarthy, Andrew. 1984. *Constraints on allomorphy in inflexion*. Bloomington, Ind: Indiana University Linguistics Club.

Carstairs-McCarthy, Andrew. 1994. Inflection classes, gender, and the principle of contrast. *Language* 70, 737-788.

Carstairs-McCarthy, Andrew. 2010. *The Evolution of Morphology.* Oxford: Oxford University Press.

Cho, Taehong. 1998. Intergestural overlap and timing in Korean palatalization. In David J. Silva (ed.), *Japanese/Korean linguistics*, vol. 8. Stanford: Stanford Linguistics Association.

Cho, Taehong. 2001. Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica* 58, 129-162.

Ciucci, Luca & Pier Marco Bertinetto. 2017. Possessive inflection in Proto-Zamucoan: A reconstruction. *Diachronica* 34(3), 283–330.

Clark, Eve. 1987. The principle of contrast: A constraint on language acquisition. In Brian MacWhinney (ed.), *Mechanisms of Language Acquisition*. Hilsdale: Lawrence Erlbaum Associates, 1-33.

Clendon, Mark. 2014. *Worrorra: a language of the north-west Kimberley Coast*. Adelaide:

University of Adelaide Press.

Conrad, Robert & Kepas Wogigia. 1991. *An Outline of Bukiyip Grammar*. (Pacific Linguistics Series C, 113). Canberra: Australian National University.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2009. Canonical inflectional classes. *Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux*, 1–11.

Corbett, Greville G. 2012. *Features*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2015. Features in inflection. In Matthew Baerman (ed.), *The Oxford handbook of inflection*. Oxford: Oxford University Press, 35–46.

Crowley, Terry. 1981. The Mpakwithi Dialect of Anguthimri. In Robert M. W. Dixon & Barry Blake (eds.), *Handbook of Australian Languages*. Amsterdam: John Benjamins, 147-196.

Cruschina, Silvio, Martin Maiden & John Charles Smith. 2013. *The Boundaries of Pure Morphology: Diachronic and Synchronic Perspectives*. Oxford: Oxford University Press.

Dahl, Osten & Maria Koptjevskaja-Tamm. 1998. Alienability splits and the grammaticalization of possessive constructions. Paper presented at the XVIth Scandinavian Conference of Linguistics.

Dammel, Antje. 2011. *Konjugationsklassenwandel. Prinzipien des Ab-, Um- und Ausbaus verbalflexivischer Allomorphie in germanischen Sprachen*. Berlin / New York: De Gruyter.

De Rijk, Rudolf P. G. 2008. *Standard Basque: A Progressive Grammar*. Cambridge, Mass: MIT Press.

Dixon, Robert M. W. 1980. *The Languages of Australia*. Cambridge: Cambridge University Press.

Donohue, Mark. 2002. Which sounds change: Descent and borrowing in the Skou family. *Oceanic Linguistics* 41 (1), 171-221.

Donohue, Mark. 2003. Agreement in the Skou language: A historical account. *Oceanic Linguistics* 42 (2), 479-498.

Donohue, Mark & Antoinette Schapper. 2008. Whence the Austronesian Indirect Possession Construction? *Oceanic Linguistics* 47 (2), 316–327.

Douglas, Wilfrid H. 1981. Watjarri. In Robert M. W. Dixon & Barry Blake (eds.), *Handbook of Australian Languages*, vol. 2. Amsterdam: Benjamins.

Eades, Diana. 1979. Gumbaynggir. In Robert M. W. Dixon & Barry Blake (eds.), *Handbook of Australian Languages*. Amsterdam: Benjamins, 245-362.

Enger, Hans-Olav. 2012. Antje Dammel. Konjugationsklassenwandel. *Beiträge zur Geschichte der Deutschen Sprache und Literatur* 134 (2), 251-258.

Enger, Hans-Olav. 2013. Vocabular Clarity meets Faroese noun declensions. *Folia Linguistica* 47 (2), 345-374.

Esher, Louise. 2012. *Future, conditional and autonomous morphology in Occitan*. DPhil thesis, University of Oxford.

Esher, Louise. 2014. Autonomous morphology and extramorphological coherence. *Morphology* 24(4), 325-350.

Esher, Louise. 2017 . Morphome death and transfiguration in the history of French. *Journal of Linguistics* 53 (1), 51-84.

Fortune, Reo. 1942. *Arapesh*. New York: Augustin.

Fouché, Pierre. 1931. *Le verbe français, étude morphologique*. Strasbourg: Publications de la Faculté des Lettres de l'Université de Strasbourg.

Franjieh, Michael. 2016. Indirect possessive hosts in North Ambrym: evidence for Gender, *Oceanic Linguistics* 55(1), 87-115.

Gardani, Francesco. 2013. *Dynamics of morphological productivity. The evolution of noun classes from Latin to Italian*. Leiden / Boston: Brill.

Gasser, Emily. 2015. The development of verbal infixation in Cenderawasih Bay. In Malcolm D. Ross & I Wayan Arka (eds.), *Language Change in Austronesian Languages: Papers from 12-ICAL*. Canberra: Asia-Pacific Linguistics.

Goddard, Cliff. 1982. Case systems and case marking in Australian languages: A new interpretation. *Australian Journal of Linguistics* 2 (2), 167-196.

Gómez, Ricardo & Koldo Sainz. 1995. On the origin of the finite forms of the Basque verb. In José Ignacio Hualde, Joseba Lakarra & Larry Trask (eds), *Towards a History of the Basque Language*. Amsterdam: Benjamins, 235-274.

Good, Jeff. 2012. How to become a "Kwa" noun. *Morphology* 22 (2), 293-335.

Gordon, Lynn. 1986. *Maricopa morphology and syntax*. Berkeley: University of California Press.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Joseph Greenberg (ed.), *Universals of human language*, vol. 3: *Word Structure*. Stanford: Stanford University Press, 47-82.

Grimm, Scott. 2010. Number and Markedness: A View from Dagaare. In Martin Prinzhorn, Viola Schmitt & Sarah Zobel (eds.), *Proceedings of Sinn und Bedeutung 14*. Vienna: s.e.

Hale, Kenneth. 1973. Deep-surface canonical disparities in relation to analysis and change: An Australian example. In Thomas Sebeok, (ed.), *Current Trends in Linguistics 11*. The Hague: Mouton, 401–458

Hale, Kenneth. 1991. Remarks on G. Sanders' "Levelling in the history of Polynesian passive formations". *Journal of the Polynesian Society* 100, 99-101.

Halle, Morris & Alec Marantz. 2008. Clarifying "blur": Paradigms, defaults, and inflectional classes. In Asaf Bachrach & Andrew Nevins (eds.), *Inflectional Identity*. Oxford: Oxford University Press, 55-73.

Hammarström, Harald, Sebastian Bank, Robert Forkel, & Martin Haspelmath. 2017. *Glottolog 3.1*. Jena: Max Planck Institute for the Science of Human History. (Available online at http://glottolog.org, Accessed on 2017-12-27.)

Harris, Martin. 1971. The verbal systems of Latin and French. *Transactions of the Philological Society* 69 (1), 62-90.

Harris, Alice & Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Harvey, Mark. 2008. The origins of conjugation markers in Australian languages. In Claire Bowern, Bethwyn Evans & Luisa Micelli (eds.), *Morphology and Language History. In honour of Harold Koch*. Amsterdam: Benjamins, 123-137.

Haspelmath, Martin. 1995. The growth of affixes in morphological reanalysis. *Yearbook of Morphology 1994*, 1-29.

Haspelmath, Martin. 2008. Creating Economical Morphosyntactic Patterns in Language Change. In Jeff Good (ed.), *Linguistic Universals and Language Change*. Oxford: Oxford University Press, 185-214.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45 (1), 31-80.

Haspelmath, Martin. 2017. Explaining alienability contrasts in adpossessive constructions: Predictability vs. iconicity. *Zeitschrift für Sprachwissenschaft* 36 (2), 193-231.

Heine, Bernd. 1973. *Pidgin-Sprachen im Bantu-Bereich*. Berlin: Reiner.

Heine, Bernd & Tania Kuteva. 2007. *The Genesis of Grammar: A Reconstruction*. Oxford: Oxford University Press.

Heuvel, Wilco van den. 2006. *Biak: description of an Austronesian language of Papua*. Utrecht: LOT.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.

Hughes, Jock. 2000. The Morphology of Dobel, Aru, with special reference to reduplication. In C. E. Grimes (ed.), *Spices from the East. Papers in Languages of Eastern Indonesia*. Canberra: Pacific Linguistics, 131-180.

Inkelas, Sharon. 2014. *The interplay of morphology and phonology*. Oxford: Oxford University Press.

Irwin, Barry. 1974. *Salt-Yui Grammar*. (Pacific Linguistics. Series B). Canberra: Australian

National University.

Jackson, Jason A. J. 2014. *A grammar of Irarutu, a language of West Papua, Indonesia, with historical analysis*. PhD thesis, University of Hawai'i at Manoa.

Karatsareas, Petros. 2011. Neuter heteroclisis in Asia Minor Greek: Origin and development. *Neoelliniki Dialektologia* [Modern Greek Dialectology] 6, 111–135.

Karubaba, Sarah. 2009. *Ambai*. MA thesis, University of Leiden

Kaye, Steven. 2015. *Conjugation class from Latin to Romance: heteroclisis in diachrony and synchrony*. DPhil thesis, University of Oxford.

Kemps, R., Ernestus, M., Schreuder, R., and Baayen, R. H. (2005). Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory and Cognition* 33, 430-446.

Kemps, R., Wurm, L., Ernestus, M., Schreuder, R., and Baayen, R. H. (2005). Prosodic cues for morphological complexity: Comparatives and agent nouns in Dutch and English. *Language and Cognitive Processes* 20, 43-73.

Kießling, Roland. 2013. On the origin of Niger-Congo nominal classification. In Ritsuko Kikusawa & Laurence A. Reid (eds.), *Historical Linguistics 2011: Selected papers from the 20th International Conference on Historical Linguistics, Osaka, 25-30 July 2011*. Amsterdam: Benjamins, 43-65.

Koefoed, Geert & Jaap van Marle. 2000. Fundamental concepts. In Christian Lehmann, Joachim Mugdan, Stavros Skopeteas, Wolfgang Kesselheim & Geert Booij (eds.), *Morphologie, 2. Halbband: Ein Internationales Handbuch zur Flexion und Wortbildung*. Berlin / New York: De Gruyter, 1574-1590.

Kristol, Andres. 2016. Francoprovençal. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press, 350-362.

Lakarra, Joseba A. 2006. Protovasco, munda y otros: Reconstrucción interna y tipología holística diacrónica. *Oihenart: Cuadernos de Lengua y Literatura* 21, 229-322.

Lass, Roger. 1984. *Phonology: an introduction to basic concepts*. Cambridge: Cambridge

University Press.

Ledgeway, Adam. 2012. *From Latin to Romance. Morphosyntactic Typology and Change*. Oxford: Oxford University Press.

Lehmann, Christian. 1995. *Thoughts on Grammaticalization*. München: Lincom Europa.

Lichtenberk, Frantisek. 1983. *A Grammar of Manam*. Honolulu: University of Hawaii Press.

Lichtenberk, Frantisek. 1985. Possessive constructions in Oceanic languages and in Proto-Oceanic. In Andrew Pawley & Lois Carrington (eds.) *Austronesian linguistics at the 15th Pacific Science congress,* 93–140..

Lichtenberk, Frantisek. 2009. Attributive possessive constructions in Oceanic. In William Mc Gregor (ed.), *The expression of possession*. Berlin: Mouton De Gruyter, 249–91.

Loporcaro, Michele. 2007. On triple auxiliation in Romance. *Linguistics* 45 (1), 173-222.

Loporcaro, Michele. 2016. Auxiliary selection and participial agreement. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press.

Luís, Ana. 2011. Morphomic structure and loan-verb integration: evidence from Lusophone creoles. In Martin Maiden, John Smith, Maria Goldbach & Olivier Hinzelin (eds.) *Morphological Autonomy: Perspectives from Romance Inflectional Morphology*. Oxford: Oxford University Press, 235-254.

Maiden, Martin. 1991. *Interactive morphonology: metaphony in Italy*. London: Routledge.

Maiden, Martin. 2005. Morphological autonomy and diachrony. *Yearbook of Morphology 2004*, 137–175.

Maiden, Martin. 2009. Where does heteroclisis come from? Evidence from Romanian dialects. *Morphology* 19 (1). 59–86.

Maiden, Martin. 2013. The Latin “third stem” and its Romance descendants. *Diachronica* 30 (4), 492–530.

Maiden, Martin. 2016a. Romanian, Istro-Romanian, Megleno-Romanian, and Aromanian. In

Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press, 91-125.

Maiden, Martin. 2016b. Number. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press, 697-707.

Maiden, Martin. 2016c. Ambiguity in Romanian word-structure. The structure of plurals in …uri. *Revue Roumaine de linguistique* 61 (1), 3-12.

Maiden, Martin. 2017. Cu privire la geneza vocalelor tematice în sistemul verbal dacoromanic. In Adina Dragomirescu, Alexandru Nicolae, Camelia Stan & Rodica Zafiu (eds.), *Sintaxa ca mod de a fi. Omagiu doamnei profesoare Gabriela Pană Dindelegan, la aniversare*. Bucureşti: Editura universităţii din Bucureşti.

Maiden, Martin. 2018. *The Romance Verb. Morphomic Structure and Diachrony.* Oxford: Oxford University Press.

Maiden, Martin, Maria Goldbach, John Charles Smith & Marc-Olivier Hinzelin. 2011. *Morphological Autonomy: Perspectives from Romance Inflectional Morphology*. Oxford: Oxford University Press.

Matthews, Peter H. 1972. *Inflectional morphology: a theoretical study based on aspects of Latin verb conjugation.* Cambridge: Cambridge University Press.

Matthews, Peter H. 1991. *Morphology*. 2nd ed. Cambridge: Cambridge University Press.

Miehe, Gudrun. 2007. Cerma (Gouin). In Gudrun Miehe & Kerstin Winkelmann (eds.), *Noun Class Systems in Gur Languages. Vol. 1. Southwestern Gur Languages (without Gurunsi)*. Köln: Rudiger Köppe.

Miehe, Gudrun. 2012. Dagara cluster. In Gudrun Miehe, Brigitte Reineke & Kerstin Winkelmann (eds.), *Noun Class Systems in Gur Languages. Vol. 2. North Central Gur Languages*. Köln: Rudiger Köppe.

Miehe, Gudrun & Kerstin Winkelmann (eds.) 2007. *Noun Class Systems in Gur Languages. Vol. 1. Southwestern Gur Languages (without Gurunsi).* Köln: Rudiger Köppe.

Milin, Petar, D.F. Durdevic & F. M. del Prado Martin. 2009. The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language* 60 (1), 50-64.

McLendon, Sally. 1975. *A grammar of Eastern Pomo*. Berkeley / London: University of California Press.

Mofu, Suriel Semuel. 2005. *Biak morphosyntax*. MPhil Thesis, University of Oxford.

Mofu, Suriel Semuel. 2008. *Biak morphosyntax*. DPhil Thesis, University of Oxford.

Mounole Hiriart-Urruty, Céline. 2006. Quelques remarques à propos de l'histoire des périphrases basques. In *R. L. Trasken Oroitzapenetan Ikerketak Euskalaritzaz eta hizkuntzalaritza historikoaz*, 723-738.

Müller, Gereon. 2007. Notes on Paradigm Economy. *Morphology* 17, 1-38.

Nichols, Johanna. 1988. On alienable and inalienable possession. In William Shipley (ed.), *In honor of Mary Haas*. De Gruyter, 557–609.

Nichols, Johanna & Balthasar Bickel. 2013. Possessive Classification. In Matthew Dryer & Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/59, Accessed on 2017-12-27.)

Norde, Muriel. 2002. The final stages of grammaticalization: affixhood and beyond. In Ilse Wischer & Gabriela Diewald (eds.), *New reflexions on grammaticalization*. Amsterdam / Philadelphia: Benjamins, 45-65.

Nordlinger, Rachel & Louisa Sadler. 2004. Nominal tense in crosslinguistic perspective. *Language* 80 (4), 776-806.

Nübling, Damaris. 2000. *Prinzipien der Irregularisierung. Eine kontrastive Analyse von zehn Verben in zehn germanischen Sprachen*. Berlin, Boston: Max Niemeyer Verlag.

O'Neill, Paul. 2011. The Notion of the Morphome. In Martin Maiden, Maria Goldbach, John Charles Smith & Marc-Olivier Hinzelin (eds.), *Morphological Autonomy: Perspectives from*

Romance Inflectional Morphology. Oxford: Oxford University Press.

O'Neill, Paul. 2014. The morphome in constructive and abstractive models of morphology. *Morphology* 24, 25–70.

Padilla-Moyano, Manuel. 2013. Non-Finite Verbal Morphology. In Mikel Martínez-Areta (ed.) *Basque and Proto-Basque. Language-Internal and Typological Approaches to Linguistic Reconstruction*. Frankfurt am Main: Peter Lang.

Parker Jones, 'Oiwi. 2008. Phonotactic Probability and the Maori Passive: A Computational Approach. *Proceedings of the Tenth Meeting of the ACL Special Interest Group on Computational Morphology and Phonology*, 39-48.

Picoche, Jacqueline. 1979. *Précis de morphologie historique du français*. Paris: Nathan.

Price, Glanville. 1971. *The French Language. Present and Past*. London: Edward Arnold.

Rice, Keren. 1989. *A grammar of Slave*. Mouton de Gruyter: Berlin.

Roberts, John Richard. 1987. *Amele*. London: Croom Helm.

Roberts, Sarah J. & Joan Bresnan. 2008. Retained inflectional morphology in pidgins: A typological study. *Linguistic Typology* 12, 269-302.

Round, Erich R. 2013. *Kayardild morphology and syntax*. Oxford: Oxford University Press.

Round, Erich R. 2015. Rhizomorphomes, meromorphomes, and metamorphomes. In Matthew Baerman, Dunstan Brown & Greville G. Corbett (eds.), *Understanding and Measuring Morphological Complexity*. Oxford: Oxford University Press, 28-52.

Round, Erich & Greville Corbett. 2017. The theory of feature systems: One feature versus two for Kayardild tense-aspect-mood. *Morphology* 27, 21-75.

Rumsey, Alan. 1982. *An intra-sentence grammar of Ungarinjin, north-western Australia*. Canberra: Australian National University.

Sanders, Gerald. 1990. On the analysis and implications of Maori verb alternations. *Lingua* 80. 149–196.

Sanders, Gerald. 1991. Levelling and reanalysis in the history of Polynesian passive formations.

Journal of the Polynesian Society, 100, 71-91.

Sawaki, Yusuf Willem. 2016. *A grammar of Wooi: An Austronesian language of Yapen Island, Western New Guinea*. PhD thesis, Canberra: Australian National University.

Schadeberg, Thilo. 2001. Number in Swahili Grammar. *Swahili Forum* 8, 7-16.

Schmid, Stephan. 2016. Segmental phonology. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press, 471-483.

Seifart, Frank. 2005. *The structure and use of shape-based noun classes in Miraña (South West Amazon)*. PhD thesis, Max Planck Institute for Psycholinguistics.

Silzer, Peter J. 1983. *Ambai: an Austronesian language of Irian Jaya, Indonesia*. Canberra: Australian National University.

Sims, Andrea D. 2015. *Inflectional Defectiveness*. Cambridge: Cambridge University Press.

Smith, John Charles. 2013. The morphome as a gradient phenomenon: evidence from Romance. In Silvio Cruschina, Martin Maiden & John Charles Smith (eds.), *The Boundaries of Pure Morphology*. Oxford: Oxford University Press, 247-261

Spencer, Andrew. 2003. Putting some order into morphology: reflections on Rice (2000) and Stump (2001). *Journal of Linguistics* 39: 1-26.

Spencer, Andrew. 2003. Periphrastic paradigms in Bulgarian. In Uwe Junghanns & Luka Szucsich (eds.), *Syntactic Structures and Morphological Information*. Berlin: Mouton De Gruyter, 249-282.

Spencer, Andrew. 2013. *Lexical relatedness : a paradigm-based model*. Oxford: Oxford University Press.

Squartini, Mario & Pier Marco Bertinetto. 1995. The simple and compound past in Romance languages. In Östen Dahl (ed.), *Tense and aspect in the languages of Europe*. Berlin: Mouton De Gruyter, 403-439.

Stebbins, Tonya N. 2011. *Mali (Baining) grammar: a language of the East New Britain province, Papua New Guinea*. Canberra: Pacific Lingustics.

Steinhauer, Hein. 2003. Konstruksi posesif dalam bahasa Biak. *Bahasa dan Sastra* 21, 1-23.

Steinhauer, Hein. 2005. Biak. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. London, New-York: Routledge, 793-823

Štichauer, Pavel. 2018. Lexical splits within periphrasis: mixed perfective auxiliation systems in Italo-Romance. *Morphology* 28 (1), 1-23.

Stump, Gregory. 2001. *Inflectional morphology : a theory of paradigm structure*. Cambridge: Cambridge University Press.

Stump, Gregory. 2006. Heteroclisis and Paradigm Linkage. *Language* 82 (2), 279-322.

Stump, Gregory. 2015. Inflection Classes. In Matthew Baerman (ed.), *The Oxford Handbook of Inflection*. Oxford: Oxford University Press.

Stump, Gregory T. 2016. *Inflectional paradigms: content and form at the syntax-morphology interface*. Cambridge: Cambridge University Press.

Stump, Gregory & Raphael Finkel. 2013. *Morphological typology : from word to paradigm*. Cambridge: Cambridge University Press.

Thomas, Amanda. 2018. *Accounting for double past participle forms in Romance*. M.Phil Thesis, University of Oxford.

Thornton, Anna. 2012. Reduction and maintenance of overabundance. A case study on Italian verb paradigms. *Word Structure* 5(2), 183–207.

Thornton, Anna. 2013. The non-canonicity of the Italian type braccio / bracci / braccia: overabundance, defectiveness, or overdifferentiation? Paper presented at the Romance Linguistics Seminar, Oxford.

Thurston, William R. 1982. *A comparative study in Anêm and Lusi*. Canberra: Australian National University.

Trask, Larry. 1995. On the History of the Non-Finite Verb Forms in Basque. In José Ignacio Hualde, Joseba Lakarra & Larry Trask (eds), *Towards a History of the Basque Language*. Amsterdam: Benjamins, 207-234.

Trask, Larry. 1997. *The History of Basque*. London / New York: Routledge.

Trommer, Jochen & Eva Zimmermann. 2015. Inflectional Exponence. In Matthew Baerman (ed.), *The Oxford Handbook of Inflection*. Oxford: Oxford University Press.

Trommer, Jochen. 2012. *The morphology and phonology of exponence*. Oxford: Oxford University Press.

Vries, Lourens de. 2004. *A Short Grammar of Inanwatan, an Endangered Language of the Bird's Head of Papua, Indonesia*. Canberra: Pacific Linguistics.

Vydrin, Valentin. 2006. Towards a reconstruction of the phonological type and the noun morphology of proto-Mande. *Acta Linguistica Petropolitana* 2, 223-246.

Walther, Géraldine. 2013. *Sur la canonicité en morphologie: perspective empirique, formelle et computationnelle*. PhD thesis, Université Paris Diderot.

Wurzel, Wolfgang. 1987. System-dependent morphological naturalness in inflection. In Wolfgang Dressler (ed.), *Leitmotifs in Natural Morphology*. Amsterdam/Philadelphia: Benjamins, 59-96.

Wurzel, Wolfgang. 1989. *Inflectional morphology and naturalness*. Dordrecht ; London: Kluwer Academic.

Zimmermann, Julia. 2016. Morphological status and acoustic realization: Is there a difference between Bra[d] Pitt and a grille[d] cheese omelet, or between Kate Mo[s] and killer robot[s]? In Johannes Wahle, Marisa Köllner, Harald Baayen, Gerhard Jäger & Tineke Baayen-Oudshoorn (eds.), *Proceedings of 6th Conference on Quantitative Investigations in Theoretical Linguistics (QITL-6), Tübingen, 4–6 November 2016*. Tübingen: Eberhard Karls Universität.

Zink, Gaston. 1986. *Morphologie du français médiéval*. Paris: PUF.