Extreme classification

Sebastian Fedden, Greville G. Corbett

Sebastian Fedden* and Greville G. Corbett

Extreme classification

https://doi.org/10.1515/cog-2017-0109
Received September 21, 2017; revised May 08, 2018; accepted May 10, 2018

Abstract: Categorization retains its key importance in research on human cognition. It is an intellectual area where all disciplines devoted to human cognition – psychology, philosophy, anthropology, and linguistics – intersect. In language, categorization is not only a central part of lexical structure but is also salient in systems of nominal classification, notably gender and classifiers. Recent years have seen great progress in the description and analysis of nominal classification systems, so that we are now in a position to offer an account of such systems which brings cognition and typology together, providing the essential parameters for the calibration of experiments for investigating cognition. To this end, we establish the extremes of nominal classification systems, from the surprisingly simple to the surprisingly complex. We analyse the two essential components of nominal classification systems: (i) assignment, i.e. the principles (semantic or formal) which govern category assignment and (ii) exponence, i.e. the morphological means by which systems of nominal classification are expressed. We discuss extreme configurations of assignment and exponence in individual languages and extreme multiple pairings of assignment and exponence in languages with two or even more concurrent classification systems.

Keywords: nominal classification, gender, classifiers, experiments in fieldwork situations

1 Introduction

Categorization continues to occupy centre stage in research on human cognition. It is an intellectual area where psychology, philosophy, anthropology, and linguistics, indeed all the disciplines devoted to human cognition, intersect. Categorization is ubiquitous: the ability to process the continuous stream of information that confronts us, to separate the signal from the noise and turn the

***Corresponding author: Sebastian Fedden,** Université Sorbonne-Nouvelle (Paris 3), France; Surrey Morphology Group, University of Surrey, Guildford, Surrey, UK(GB),
E-mail: sebastian.fedden@univ-paris3.fr
Greville G. Corbett, Surrey Morphology Group, University of Surrey, Guildford, Surrey, UK(GB),
E-mail: g.corbett@surrey.ac.uk

former into manageable units, is crucial for dealing with the world around us and especially with our fellow human beings (see Koestler 1983: 201, cited in Senft 2007: 676). Establishing meaningful units – categorization – is thus a process fundamental to thought and communication. And the way we do this reveals interesting differences across languages and cultures, in that the same real-world entities may be treated very differently. For example, at the lexical level, speakers of English differentiate between *fingers* and *toes*, while for speakers of Italian they are all referred to by the same word, *dita*.

Categorization is not only an important part of lexical structure, it is particularly salient in systems of nominal classification (Senft 2007), since it is here that grammatical rules of a language can force speakers to classify. The pronoun system of Standard English requires us to choose between *he* for males, *she* for females and *it* for inanimates along the lines of biological sex. In a language with a gender system like Italian, all nouns are treated as either masculine or feminine. Human nouns are usually assigned their gender according to their semantics: *uomo* 'man' is masculine and *donna* 'woman' feminine', but even those nouns whose meanings have nothing to do with biological sex need to be part of the gender system and are thus classified as either masculine, e.g. *sasso* 'rock, stone', or feminine, e.g. *pietra* 'rock, stone'. There has been careful recent discussion of the relations between linguistic gender, culture, and cognition in Beller et al. (2015) and Bender et al. (2016); see also the references in both. For work on classifiers, suggesting that they reflect conceptual structure, rather than affect it, see Speed et al. (2016).

While the familiar languages with smaller systems of nominal classification have attracted the most interest, there are other languages with much more elaborate systems of categorization. Thus the Oceanic language Kilivila (Senft 1993, 1996), spoken on the Trobriand Islands in Papua New Guinea, has no fewer than 177 distinct possibilities that speakers of the language have to choose from to produce the appropriate forms of numerals, a subset of adjectives, and most demonstratives. In Kilivila, categorization is based on fine-grained meaning contrasts, including shape, function, arrangement, place or time interval. Traditionally, languages like Kilivila, along with many languages of east and southeast Asia, have been analyzed as "classifier" languages, and typologists have devised a set of criteria attempting to establish a clear opposition between "gender" systems and "classifier" systems (Dixon 1982, 1986). But since the inception of this opposition there has been doubt about its validity (cf. Gomez-Imbert 1982). More recent language data and analysis show that we have to rethink the old typologies (Corbett and Fedden 2016; Fedden and Corbett 2017a, 2017b). For example, the Witotoan language Bora-Miraña (Seifart 2005) has a system of nominal classification which is an intermediate one; it combines

properties of what typologists traditionally called "gender" and what they traditionally called "classifiers".

As typologists, we are interested in the limits on language variation. This makes the extremes of classification a natural area of study. But extremes are also vital for the calibration of our experimental methods for investigating cognition. As equipment gets cheaper and more portable, it is becoming feasible to carry out sophisticated experiments in field situations (Harris and Samuel 2011; Harris 2017: 91–95). Techniques such as eye-tracking allow us to ask questions about the relation between language and cognition which until recently could not be investigated. And as the range of languages available for this type of work expands, we need to ensure that experimenters' choice of languages includes the extremes. One goal of our paper, therefore, in laying out the extremes of categorization, is to facilitate experimental design. Thus where we can demonstrate that systems of nominal classification vary along some dimension, we shall plot the extremes. As an example (and previewing Section 2.1), consider how nouns are assigned a gender. At one extreme, the generalizations needed are few (as in the Nakh-Dagestanian language Bagvalal, Section 2.1.1). At the other extreme are languages like German (Section 2.1.2), where gender assignment requires a large set of descriptive statements. Both are interesting from a cognitive perspective: given how simple a system of categorization can be, as in Bagvalal, it is striking that some languages go to the other extreme, as German does. And we would not wish for our experiments to include only one of these extremes (except by deliberate and motivated choice). It is also striking that there are languages which combine these two extremes. An example is Archi (Section 2.1.3), whose assignment system has elements partly reminiscent of German and of Bagvalal.

Audring (2014, 2017) measures the complexity of gender systems along various dimensions, among them the number of gender values and the number and nature of assignment rules; the basic idea behind this is that more gender values and more complicated assignment rules *ceteris paribus* make for more complex gender systems (see also Enger 2011).[1] The notion of 'extreme' that we use in this paper covers both highly complex and highly simple systems and will also be applied to the various situations of concurrency where a language has more than one system of the same type.

1 There is a substantial literature on complexity; for work specifically concerned with gender, see Di Garbo and Wälchli (Forthcoming), and for complexity more generally see, for instance, Dahl (2004) and Baechler and Seiler (2016).

In probing the extremes, we come to the heart of the issue of categorization. On the one hand, an effective categorization system has to be simple, in order to minimize cognitive load. On the other hand, it must be informative, to maximize communicative efficiency. There is a trade-off between the principles of simplicity and informativeness (Hawkins 2004). The extensive variation we document in the nominal classification systems of the world's languages results from the competition between the two principles. Rosch suggested that categorization systems should "provide maximum information with the least cognitive effort" (1978: 190). We can start to ask how the extreme systems we document square with this requirement. We can begin to investigate how information is stored and retrieved, in terms of measurable factors such as speed, accuracy and inter-speaker consistency.

In the following sections, we discuss the extremes of nominal classification systems. In Section 2 we discuss the essentials: categorization (assignment), form (exponence), and their cognitive connection. Section 3 is devoted to extremely simple systems, particularly that of Bagvalal. In Section 4 we move on to more complex relations of assignment and exponence. In Section 5 we deal with concurrent systems, i.e. languages with two classification systems. Section 6 is about extreme concurrency and those cases for which more than two concurrent systems have been claimed to exist.

2 The essentials: Categorization (assignment) and form (exponence) and their extremes

There are two essential parts of a system of classification: the distinctions and the outcome in form. From the speaker's perspective, a value must be selected within the system of classification (in many systems this selection may be completely determined), and this selection must be realized in linguistic form. Models of the first part are termed *assignment systems*; in Italian, as mentioned above, we say that *uomo* 'man' is assigned masculine gender, and *donna* 'woman' is assigned feminine gender. These values are reflected in the different agreement targets, that is, through the system of exponence of gender. Thus assignment and exponence go hand in hand. Exponence is the morphological means by which systems of nominal classification are expressed: the evidence for a system of classification is precisely the systems of exponence. The system of exponence requires that nouns have a gender value assigned to them, and without exponence we could not analyse the system of assignment.

2.1 Assignment

For assignment, we can imagine two extremes: (i) a semantically fully transparent system, where we can reliably predict the category of a noun from its meaning and (ii) a semantically fully opaque or arbitrary system, in which category membership is completely random. In reality these are not on equal footing. Semantically fully transparent assignment systems actually exist (as we shall see in Section 2.1.1), whereas fully random ones do not. Categorization in all nominal classification systems proceeds according to semantic principles, at least for a subset of nouns. This semantic core typically contains humans and higher animals (Corbett 1991: 7–32).[2] But what we do find are languages which beyond the semantic core have an intricate system of phonological and morphological assignment rules, some of which only allow us to predict the category of a noun with a certain probability.

2.1.1 Extreme 1: Fully transparent systems

A semantically fully transparent assignment system can be found in the Nakh-Dagestanian language Bagvalal (Kibrik et al. 2001), spoken in southwestern Dagestan by approximately 1,500 speakers. In Bagvalal, there is evidence from agreement for three gender values. Assignment is fully semantic: nouns denoting male humans are masculine; nouns denoting female humans are feminine; all remaining nouns are neuter. The neuter gender comprises all non-humans (whether animate or inanimate):

Bagvalal

(1) *waša* *w-iRi*
boy M.SG-stop
'the boy stopped' (Kibrik et al. 2001: 64)

(2) *jaš* *j-iRi*
girl F.SG-stop
'the girl stopped' (Kibrik et al. 2001: 64)

2 Motivation wins out: where there is a clash, the major semantic assignment rules take precedence over formal assignment (Corbett 1991: 33–69); for discussion see Nesset (2006), Enger (2009) and Thornton (2009). Interestingly, however, various studies suggest that children tend to give undue weight to phonological cues early in acquisition of gender systems. A language claimed to have a gender system with a high degree of opacity is Uduk (Killian 2015: 67–73), but the data are still limited. For recent discussion, based on statistical models, see Gagliardi et al. (2017).

(3) *ʕama b-iRi*
donkey N.SG-stop
'the donkey stopped' (Kibrik et al. 2001: 65)

Similar systems are also known from the Dravidian family, for example Tamil (Asher 1989 [1985]: 36–37; Corbett 1991: 8–9). English also fits here, for those who accept gender systems with only anaphoric pronouns as agreement targets. Fully semantic assignment holds for the standard language, but see Pawley (2002) and Siemund (2008) for the use of *he* or *she* for inanimates in varieties of English.

We might wonder whether there is an even simpler solution of organizing gender assignment. A gender system with just two values would be simpler; it requires fewer distinctions. But in two-value gender systems based on sex we typically encounter other complications. In Bagvalal there is a simple mapping in both directions: male humans are of masculine gender and all nouns of masculine gender denote male humans. In a two-gender system, there is the issue of what to do with the nouns below the threshold of sex-differentiability (lower animals and inanimates). One possibility is to collapse the neuter (as found in Bagvalal) with one of the human genders. The Dravidian language Parji (Burrow and Bhattacharya 1953: 9) has two genders, one for male humans and one for the semantic residue. Another possibility is to go one step further and lower the threshold for sex-differentiability. This is what we find in the Australian language Diyari (Austin 2013 [1981]: 60), which has one gender for females, including not only humans but also higher animals, such as kangaroos and dogs, and one gender for the semantic residue. Compared to Bagvalal these are simpler systems because they only involve two gender values, but the mapping in Parji and Diyari is more involved than in Bagvalal. While Bagvalal allows prediction in both directions, semantics to gender and gender to semantics, in Parji and Diyari we often cannot predict the semantics from the gender value. In Parji we can predict the semantics only if the noun is masculine; if it is non-masculine it could be anything but a male human. Mapping in Diyari is even more involved. Feminine nouns denote female beings, and for non-feminine nouns all prediction fails because they can have a wide range of meanings from human males to inanimates.

A third possibility is not to use biological sex as an assignment principle in the gender system and instead rely on a distinction based on animacy. This is what we find in Algonquian languages, where the genders are animate and inanimate. But even these systems are rarely simple, in that we typically get various 'leaks' of notionally inanimate nouns into the animate gender, some of which are famous (like 'raspberry'). However, a more consistent two-valued

gender system has been identified, namely Sumerian (Passer 2016: 221, 564–571). Sumerian was a language isolate of southern Mesopotamia. It had two gender values, human and non-human, with highly consistent assignment to them and few deviations (Jagersma 2010: 101–105).

Semantically fully transparent assignment systems are one extreme of classification. A clear example is the three-gender system of Bagvalal. While the two-gender systems in this section are simpler in numbers of values they are not automatically simpler in terms of assignment (see also Audring 2017 on this distinction).

2.1.2 Extreme 2: Large set of assignment rules

German is a good example of a language with a notoriously intricate assignment system. German has masculine, feminine and neuter genders. Some nouns are assigned their gender based on the semantics, for example, *Mann* 'man' and *Frau* 'woman' are masculine and feminine, respectively. However, German also has many inanimate masculine nouns, e.g. *Kamm* 'comb', *Schuh* 'shoe'. For these nouns gender cannot be read off the semantics, since they lack biological sex and therefore – from a semantic perspective – should be neuter. This is where formal assignment rules come in. The form of these nouns points towards masculine gender, given that most monosyllabic nouns in German are masculine (see Köpcke and Zubin 1984: 29), and the fuller inflectional paradigm is also a strong predictor. In fact, assignment in German can get rather involved. Köpcke (1982: 69–108) develops 44 phonological, morphological and semantic rules which account for the gender assignment of 90% of monosyllabic German nouns. The interest of German as an example of an extreme assignment system lies not only in the fact that a large number of assignment rules is required, but also that these rules have a substantial number of exceptions. In many cases the form helps (i.e. the phonology or the morphology), but often only statistically. For example, we can predict that *Knauf* 'knob' (masculine) cannot be feminine based on its paradigm, and we know that it is very likely masculine, given that almost all monosyllabic German nouns starting with the consonant cluster /kn/ are masculine, the only exception being the neuter noun *Knie* 'knee' (Köpcke and Zubin 1984: 29–30).

Languages like German raise fascinating questions in terms of gender assignment. They involve what Koenig (1999: 1–2) has nicely termed "medium-size generalizations". There are interrelated generalizations, with varying degrees both of successful coverage of the data and of cognitive plausibility. For discussion see, among others, Corbett (1991: 7–69, 2014: 110–124),

Evans et al. (2002), Nesset (2006), Enger (2009), Thornton (2009), and Plaster and Polinsky (2010).[3]

2.1.3 A combination of extremes 1 and 2: Archi

Bagvalal and German are extremes of assignment systems. Bagvalal is simple and requires few generalizations. German is more complex, in that each gender value includes nouns which belong there for a good semantic reason, and others which do not. The number of generalizations involved is much higher than in Bagvalal. Now we turn to the interesting case of Archi, another language from Dagestan, which combines properties of Bagvalal and at least to some extent properties of German.

The gender and number agreement system of Archi can be represented as in Table 1 (*x-* is the prefixal form, and ‹*x*› the infixal form); the original source is Kibrik et al. (1977: 55–66).

Table 1: Gender and number in Archi (verbal agreement).

GENDER	Assignment	NUMBER	
		SINGULAR	PLURAL
I	male human	w-/‹w›	b-/‹b›
II	female human	d-/‹r›	
III	some animates, all insects, some inanimates	b-/‹b›	Ø-/‹Ø›
IV	some animates, some inanimates, abstracts	Ø-/‹Ø›	

In Archi there is a four-way distinction of gender values in the singular. In the plural these collapse to two. While the forms themselves are interesting and challenging (Chumakina and Corbett 2015) our focus here is on gender assignment. The assignment of gender values I and II is semantic and straightforward. For gender values III and IV, despite their relations to semantic categories, the assignment is not straightforward. Given an inanimate object of unremarkable size there is (as yet) no clear prediction of its gender.

3 For the reaction of such systems to attrition, see Lohndal and Westergaard (2016), and references there; for second language acquisition, see Binanzer (2017).

This means that Archi has two gender values, I and II, which are semantically transparent, very similar to Bagvalal. And then it has two more, III and IV, which are less straightforward; while not as complex as those of German, they are certainly closer to German genders than I and II are. This interesting combination, transparency in one part of the systems and a degree of opacity in another, marks another type of extreme. Rather than simply fitting within the dimension fewer–more generalizations for assignment, Archi differs in splitting its gender values, and having a combination of two different types of assignment system. (We return to the issue of motivation in Section 3.)

2.2 Exponence

Nominal classification systems can be extreme in terms of the 'visibility' of their exponents, that is, in terms of the evidence there is for gender in the language. This varies from extreme visibility to minimal visibility, which has implications both for function[4] and for language acquisition (see for example Audring 2014; Gagliardi and Lidz 2014). For the first type of extreme we will stay with Archi. This language has an extreme agreement system, in which almost all parts of speech agree (Chumakina and Corbett 2008; Bond et al. 2016). The agreement system and therefore the classifications which it encodes are thus highly visible. Targets in Archi agree in gender (I–IV) and number (singular or plural), using the forms given earlier in Table 1. This is illustrated in (4), with the agreement markers indicated in bold, to give an impression of the visibility of gender:

(4) Archi

*nena‹**b**›u*	[*do:ˤzu-**b***	*χˤon*]$_{NP}$	
1PL.INCL.ERG‹III.SG›	be.big.ATTR-III.SG	cow(III)[SG.ABS]	
***b**-ela‹**b**›u*	*dit:a‹**b**›u*	*χir*	*a‹**b**›u*
III.SG-1PL.INCL.DAT‹III.SG›	quickly‹III.SG›	behind	‹III.SG›make.PFV

'We quickly drove the big cow to us (home).' (Bond et al. 2016: 3)

Archi has an ergative-absolutive agreement system. In (4), the agreement controller is the absolutive argument *χˤon* 'cow', a gender III noun in the singular. As one might expect, the attributive modifier *do:ˤzub* 'big' agrees in gender and number with its head noun within the domain of the NP. In addition, we find agreement marked on a wide range of other targets within the clause. The main

4 See Acuña-Fariña (2016) for a review of psycholinguistic work on agreement, including agreement in gender.

verb *abu* 'make' at the end of the clause agrees, but so too do the ergative subject *nenabu* 'we (INCL)' and the adverb *dit:abu* 'quickly', which are typologically much rarer agreement targets.[5] Note particularly the indirect object pronoun ***b**ela**b**u* 'to us (INCL)', which not only marks gender and number agreement, but does so twice. This is an instance of 'multiple exponence' (for which see the fine recent discussion in Harris 2017). Multiple exponence makes the gender system yet more visible. Archi represents another extreme here, with some examples of multiple exponence involving four markers of gender-number (see Kibrik 1977: 128–130, 320 discussed in Corbett 1991: 108 and Harris 2017: 204).

The logical opposite extreme of a system with a wealth of agreement targets like Archi is a system where only a single part of speech agrees. We can find this situation in North Ambrym, an Oceanic language of Central Vanuatu. North Ambrym is analysed as having possessive classifiers of the type *ye-* 'edible' in Franjieh (2012, 2018), as in (5), and *ma-* 'drinkable', as in (6).

(5) North Ambrym
ye-ng *barrbarr*
CLF:edible-1SG pig
'my pig' (Franjieh 2018: 40)

(6) *ma-n* *we*
CLF:drinkable-3SG water
'his/her water' (Franjieh 2018: 37)

Such classifiers are typically found in Oceanic languages and are often called relational classifiers, where the classifier is said to characterize the relation between the possessor and an (alienably) possessed object, according to the intended use of the possessed by the possessor (Lichtenberk 1983). Using psycholinguistic experiments, Franjieh (2012) shows for North Ambrym that in many cases the classifier does not change with intended use. So examples (5) and (6) above are possible, and in fact required, if one wants to talk about pigs and water outside of the context of ingestion. This makes the categories in North Ambrym look more like the values of a morphosyntactic feature and the classifiers more like agreement markers on possessive pronouns in the context of alienable possession.

5 Note, however, that while in Archi almost all parts of speech agree, large numbers of lexical items do not. For instance, there are verbs which agree, but many do not. See Chumakina and Bond (2016: 111–116) for information, including some statistics.

North Ambrym is an example of an extreme agreement system not only because of the dearth of agreement targets – there is only one – but also because agreement is restricted further, namely to the context of alienable possession.

2.3 Cognitive connection between assignment and exponence

At first glance, these two components of a gender system – assignment and exponence – can be configured in different ways, and it might seem that all combinations are possible. Bagvalal has a simple semantic assignment system and extensive exponence. Languages like the Slavonic language Russian (Corbett 1991: 34–43) and the Bantu language Chichewa (Corbett and Mtenje 1987), as examples of a common pattern, show more complicated systems of assignment, involving formal as well as semantic rules, and both have extensive exponence.[6] However, there is an interesting regularity here, as pointed out by Audring (2014: 14). Languages like English, Diyari, and Malayalam have very restricted exponence of gender, limited in fact to the pronouns. Such languages always have simple, semantically-based assignment systems. Put another way, complex assignment systems, involving rules going beyond the basic semantic type, require substantial exponence to maintain them.[7] Once this regularity is pointed out, it makes good sense: in order to learn the more complex type of system, the child requires more evidence, in the shape of inflectional morphology, than is provided by languages of the English type. To sum up, the attested configurations of the essential components of a gender system are at least partially constrained by cognition.

3 Extreme simplicity: A system close to canonical

So far we have operated with an intuitive notion of 'extreme'. As more criteria come into play, we shall need a cleaner notion of extreme. When we find variation along some dimension, we can typically anchor one end of the scale. Just as we measure length from zero, so we can anchor linguistic measures. This is where the idea of a canon is valuable: we set up a canonical or ideal point,

6 Larger systems of values, as in Chichewa and other Bantu languages, might appear challenging for learners. Data from the acquisition of the Bantu language Sesotho, which also has pervasive agreement, suggest the contrary (Demuth 1988). Faced with extensive evidence for the system, learners acquire it early and "error free" (Demuth 2003: 213).

7 And see Rădulescu and Beuls (2016) for modelling of the development of this type of simple assignment system in Dutch, and Kraaikamp (2017) for the textual evidence.

and use it as a baseline to calibrate from (see, for example, Corbett 2012, 2015; Bond 2013; Brown and Chumakina 2013; Nikolaeva 2013; Michael 2014; Kwon and Round 2015; Forker 2016; Stump 2016: 31–42; Evans et al. 2018).

Specifically for nominal classification, we set up canonical gender as our baseline. Above we have introduced the Nakh-Dagestanian language Bagvalal as an example of an extreme – that is extremely simple – assignment system. In this section we will look at Bagvalal more closely as a language which has a gender system close to canonical in the sense of Corbett and Fedden (2016). They define canonical gender using three criteria which fall under the Canonical Gender Principle (CGP): IN A CANONICAL GENDER SYSTEM, EACH NOUN HAS A SINGLE GENDER VALUE. In the following we go through the three criteria and show how Bagvalal behaves with respect to each of them.

(i) Canonical Gender – Criterion 1 (Corbett and Fedden 2016: 505)
Canonical gender values match agreement classes.

The recognized analytical technique for establishing the number of genders in a language is to establish agreement classes based on syntactic evidence. This approach goes back to Zalizniak (1964); see also Corbett (2012: 80–85) and Mel´Čuk (2013). The idea is that nouns are in the same agreement class provided that given the same conditions they will control the same agreement form.

Typically we recognize fewer genders than agreement classes because languages have inquorate genders or subgenders. While all of these increase the number of agreement classes, counting each additional agreement class as its own gender is rarely warranted. Bagvalal is almost fully canonical with respect to the first criterion for canonical gender. The number of agreement classes equals the number of genders: there are three agreement classes and three genders in Bagvalal. The exception is a small number of nouns denoting humans which are not specific as to sex.[8]

We now move on to the second criterion for canonical gender.

(ii) Canonical Gender – Criterion 2 (Corbett and Fedden 2016: 517)
In a canonical gender system the gender of a noun is constant across all domains in which a given language shows agreement.

8 To present these, we must first explain that, in the plural, Bagvalal has syncretism between the masculine and feminine, so that there is a human versus non-human distinction in the plural. The examples which fall outside the three main agreement classes include the non-sex specific *aram* 'person' and *mač*' 'baby', which are neuter when singular, and in the plural take this syncretic masculine and feminine (human) agreement (Kibrik et al. 2001: 458). When denoting a male, each of these takes masculine agreement (see Kibrik et al. 2001: 458 and, 2001: 796, Text 7, sentence 8 for a textual example).

Gender agreement can manifest itself across different domains. Closest to the controller we find agreement in the noun phrase. The predicate can show agreement within the clause and free pronouns agree anaphorically across clause and sentence boundaries. The second criterion for canonical gender requires consistent agreement across domains. A well-known source of non-canonicity in this regard are lexical hybrids, which allow different agreement patterns in different domains, for example the German noun *Mädchen* 'girl' requires neuter (i.e. syntactic) agreement in adjectives and determiners within the noun phrase, but allows either neuter agreement or feminine (i.e. semantic) agreement in the free pronoun. Bagvalal again is close to canonical in this respect. It has few lexical hybrids.[9]

Finally, we come to the third criterion for canonical gender.

(iii) Canonical Gender – Criterion 3 (Corbett and Fedden 2016: 520)
In a canonical gender assignment system, the gender of a noun can be read unambiguously off its lexical entry.

This criterion has already figured prominently in Section 2.1.1. Bagvalal has a strictly semantic assignment system. For each noun we can predict the gender based on its meaning: nouns denoting male humans are masculine, nouns denoting female humans are feminine, and all remaining nouns are neuter. Therefore, Bagvaval is canonical with respect to the third criterion; the potential issues discussed earlier in footnotes (nouns denoting humans without respect to sex, and nouns denoting collections of humans) are all at least partly semantically predictable.

We have taken full motivation as our extreme, the canonical point. But we should recall Taylor's important point (1989: viii) that categories are motivated "to varying degrees". We look for motivation and respect the data where the evidence points against it. Thus the famous example of gender in Dyirbal, described in detail by Dixon (1972: 44–47, 60–62, 306–312, 1982: 178–183), and made famous in Lakoff (1987: 91–104), has been reanalysed by Plaster and

9 These are nouns denoting groups of people, like *ahlo* 'people' (and again *aram* 'person, people'), which can take neuter singular and masculine/feminine (human) plural agreements. See Kibrik et al. (2001: 484–485, and 480 for a further textual example). Two things should be noted here. First, in these instances the issue is primarily one of number, and the gender difference follows from the earlier point about the neuter being used in the singular when the sex of the referent is not specific. And second, the data available, limited to the noun phrase and the clause (Kibrik et al. 2001: 485), are in accord with the Agreement Hierarchy (Corbett 2006: 206–230).

Polinsky (2010) in a way that is simpler and more plausible, when factors other than semantic motivation are taken carefully into account.

An example much further from canonical for this criterion would be German, which has also been treated in some detail above (Section 2.1.2).

Bagvalal is canonical with respect to all three criteria for canonical gender. Its genders match its agreement classes (Criterion 1), agreement patterns are consistent across domains (Criterion 2), and the assignment system is strictly semantic (Criterion 3). Given this, we can say that Bagvalal observes the canonical gender Principle, according to which each noun has a single gender value. However, while canonical, even Bagvalal presents a few slight wrinkles. There are some human nouns which are not sex-specific and there is a systematic syncretism of the masculine and feminine forms in the plural thus neutralizing the contrast between the two, yielding a human vs. non-human contrast.

It has proved useful to set up a canonical baseline to calibrate from. While we do not expect to find exemplars fully matching such a baseline, it is striking just how close to it Bagvalal comes. And, given that such simplicity is attested, it is all the more remarkable that other extant systems are a long way from this simplicity.

4 More complex relations of categorization and exponence

We now turn to situations where the relations between categorization and exponence are more complex. We begin with Nyan'gityemerri (Section 4.1), which arguably has a single gender system, but shows extreme variation in the means of exponence for different values. This peculiar situation arises from the incipient nature of gender in the language. Bora-Miraña (Section 4.2), on the other hand, has one means of exponence – there is one slot in the agreement targets – but there is more than one means of categorization which have to collapse onto each other because in terms of exponence they are competing for the same slot.

4.1 Ngan'gityemerri

The Australian language Ngan'gityemerri (Non-Pama-Nyungan, Southern Daly; Reid 1997) offers us a fascinating view of how gender systems develop. It shows

how generic nouns (classifiers) which appeared in front of nouns but also before modifiers became free agreement forms with these modifiers, and then bound agreement forms.

Looking first at the semantics, Ngan'gityemerri's classification system distinguishes 15 genders, including male, female, canine, animal, vegetable, tree/thing, but also has semantically very specific genders for long woomeras, canegrass spears and digging sticks. Assignment is predominantly semantic with some leakage (Reid 1997: 165), e.g. the animal gender includes a few body parts, as well as corpses, spiritual beings, money and playing cards, and the tree/thing gender contains songs.

The formal realizations are of considerable interest. The agreements are highly non-canonical, and varied. Since the variation in type of marker partly matches the forms appearing on the noun itself to express gender overtly, we examine these for comparison. There are agreements on these targets: adjectives, demonstratives, numerals and possessive pronouns. The genders can be grouped depending on the morphological status of their forms. For six genders marking of the noun and agreement marking are done by free forms. The rest of the genders express agreement by a bound form. These can be further split according to whether the overt gender marker on the noun is a proclitic or a prefix. The bound forms are contracted from the free forms in a relatively transparent fashion. Table 2 lists all the gender markers and the gender marking patterns.

Table 2: Gender markers and marking patterns in Ngan'gityemerri (based on Reid 1997: 173).

Gender marking pattern		Gender	Noun marking		Agreement marking	
			Free	Bound	Free	Bound
Both overt gender and agreement marked by proclitic	1	male	(*mipurr*)[10]	*wa=*		*wa=*
	2	female	(*falmi*)	*wur=*		*wur=*
	3	group		*awa=*		*awa=*

(*continued*)

10 *Mipurr* and *falmi* mean 'man' and 'woman', respectively. They appear in brackets because they do not display all properties of generic nouns.

Table 2: (*continued*)

Gender marking pattern		Gender	Noun marking		Agreement marking	
			Free	Bound	Free	Bound
Overt gender marked by prefix, agreement marked by proclitic	4	animal	*gagu*	*a-*		*a=*
	5	vegetable	*miyi*	*mi-*		*mi=~ yerr=*
	6	body parts[11]		*da- ~ a-*		*a=*
	7	canines		*wu-*		*wu=*
	8	tree, thing	*yawurr*	*yerr-*		*yerr=*
	9	bamboo spear	*yawul*	*Ø-*		*yeli=*
Both overt gender and agreement marked by free form	10	fire	*yenggi*		*yenggi*	
	11	strikers	*syiri*		*syiri*	
	12	canegrass spear	*kurum*		*kurum*	
	13	drinks	*kuru*		*kuru*	
	14	long woomeras	*tyin*		*tyin*	
	15	digging sticks	*kini*		*kini*	

While bound gender markers on the noun are obligatory for nouns that have them,[12] agreement marking by means of a clitic or a free form is optional (Reid 1997: 168). This makes Ngan'gityemerri non-canonical in this respect (Corbett 2006: 14; Corbett and Fedden 2016).

As we see in Table 2, there are three distinct patterns of gender marking. First, for genders 1–3, the gender marker on both the head noun and the modifier can be a proclitic (marked by=), as in (7).

11 This is a class defined only by overt marking on the noun ('head marking' in Reid's terms). According to their agreements, a few body parts are in the animal gender, the rest belongs to the unmarked residue gender. We retain the body part class to conform to Reid's table.

12 For most gender values there are some nouns which take the overt marker and some which do not, e. g., *wamanggal* 'wallaby', not **a-wamanggal*; we know that *wamanggal* belongs to the animal gender because it takes *a=* agreement (Reid 1997: 173).

(7) Ngan'gityemerri
wa=tyerrmusye (*wa=*)*mirrisyarra* *perrety-meny*
MALE=old.man (MALE=)blind die-3SG.SBJ:do
'The old blind man has died.' (Reid 1997: 174)

Second, for genders 10 to 15 at the bottom of Table 2, generic nouns can optionally precede specific nouns and modifiers, as in (8).

(8) (*kurum*) *yiliyili* (*kurum*) *ngayi kide*
(CANEGRASS) mangrove.tipped.spear (CANEGRASS) mine where
'Where is my mangrove-tipped *kurum* spear?' (Reid 1997: 177)

The first instance of *kurum*, which is used for canegrass spears, has the function of a generic noun, whereas the second instance is more like a (free form) marker of agreement.

Third, for six genders (genders 4 to 9) there is an intermediate status: the overt gender marker is a prefix and the agreements are proclitics. This pattern is illustrated in (9).

(9) *a-matyi* (*a=*)*minbadi*
ANIMAL-kangaroo ANIMAL=big
'a big kangaroo' (Reid 1997: 181)

With four of these six genders a free generic noun can optionally be present, as in (10).

(10) (*gagu*) *a-matyi* *bengin-da*
(animal) ANIMAL-kangaroo 3SG.SBJ:AUXILIARY-hit
'He shot a kangaroo.' (Reid 1997: 175)

In the development of Ngan'gityemerri, generic nouns were funneled into the gender system by becoming free form agreements distributed throughout the noun phrase, and then bound (clitic) agreements (Reid 1997: 218). In the expression of overt gender on nouns this went even further in that the overt gender markers became prefixes.[13] So while the Ngan'gityemerri system shows considerable variation in its formal realization due to its historical development, when we look at the semantics of the system we see a clear single system.

13 Reid (1997: 212–215) gives evidence from stress placement and assimilation to distinguish between clitics and affixes.

Ngan'gityemerri is extreme in terms of its means of exponence due to the highly varied nature of agreement marking.

4.2 Bora-Miraña

Bora-Miraña is a Witotoan language spoken by about 2,400 people in Peru and by a few hundred in Colombia (Grinevald and Seifart 2004; Seifart 2005, 2009). The language uses multiple means of categorization, in the sense that there are two sets of markers controlled by nouns, which Seifart labels *specific class markers* and *general class markers*. The first set consists of 66 specific class markers (SCMs), and these are available only for lower animals and inanimates. For the full set of forms, see Seifart (2005: 86–94). A noun with a SCM is interpreted as singular, and then dual and plural can be expressed agglutinatively by means of a further suffix. The SCMs form an obligatory agreement system. Here is an example (11).

(11) Bora-Miraña
ó-di *íhka-ko* *tsa-ko*
1SG-POSS COP-SCM.1D.POINTED one-SCM.1D.POINTED
pihú-ko
fishing.rod-SCM.1D.POINTED
'I have one fishing rod.' (lit. one fishing rod is mine) (Seifart 2005: 5)

The second, much smaller set consists of general class markers (GCMs), which make distinctions along the lines of animacy and sex and cumulatively also express number. All the GMCs are given in Table 3.

Table 3: Bora-Miraña GCMs (adapted from Seifart 2005: 83)[14].

	SINGULAR	DUAL	PLURAL
MASCULINE	*-ːbɛ*	*-mɯtsi*	*-mɛ*
FEMININE	*-dʒɛ*	*-mɯpɨ*	
INANIMATE		*-nɛ*	

14 We omit the various allomorphs of these forms.

The GCMs occur on exactly the same range of targets as the SCMs.[15] Agreement targets are various types of pronoun (personal, demonstrative, possessive, and interrogative), numerals, relative clauses, and main clause predicates.

Bora-Miraña fits into our discussion of extreme nominal classification systems because it uses different markers on the same targets. This is unexpected. Rather we would expect to find the same (or almost the same) markers on different targets, and indeed there are such languages (Lamnso and Kilivila are examples, discussed in detail in Fedden and Corbett 2017a: 20–21). A Kilivila example is provided in (12).

(12) Kilivila (Senft 1986: 69)

mi-na-si-na	*na-yu*	*na-manabweta*	*vivila*
this-FEMALE-PL-this	FEMALE-two	FEMALE-beautiful	girl

'these two beautiful girls'

Else we may find different markers on different targets. Some familiar languages behave this way. For example, in French, feminine gender in the singular is realized differently on the article, *la*, and the adjective (final consonant depending on the adjective, e.g. *blanche* /blɑ̃-ʃ/ [white-F.SG]; *contente* /kõtɑ̃-t/ [happy-F.SG]).

Bora-Miraña confounds our reasonable expectations in that we find different markers on the same targets. Nouns denoting humans and higher animals require GCM agreements, but the Bora-Miraña system allows GCMs to be used as agreements with any noun, even inanimates (Seifart 2005: 79–80). Example (13) is identical to example (11), except that the GCM *-nɛ* for inanimates is used on the copula and the numeral instead of the semantically more specific SCM *-ko* '1D pointed', which is employed in (11). This alternative marking only applies to agreement, the SCM *-ko* '1D pointed' on the noun itself does not change.

(13) Bora-Miraña

ó-di	*íhka-nɛ*	*tsa-nɛ*	*pihú-ko*
1SG-POSS	COP-GCM.INAN	one-GCM.INAN	fishing.rod-SCM.1D.POINTED

'I have one fishing rod.' (lit. one fishing rod is mine) (Seifart 2005: 5)

15 Since there are alternative agreements on the same targets, SCMs versus GCMs, we might expect there to be Agreement Hierarchy effects. Seifart (2018) discusses this in a diachronic context.

The use of SCMs, on the other hand, is restricted to lower animals and inanimates. It represents a more finely-grained classification of the inanimates. Nouns lacking an SCM do not 'escape' as unclassified: agreement is obligatory so that those nouns for which there is no SCM (the animates) must take the appropriate GCM. Due to the competition created by the fact that targets have only one agreement slot the two sets of markers have to collapse onto each other. The mapping between the two sets of markers for inanimates is many-to-one, as shown in Table 4.

Table 4: Mapping between GCMs and SCMs in Bora-Miraña.

	GCMs	SCMs
Human and higher animals	MASCULINE / FEMININE	none: GCM is used
Lower animals and inanimates	INANIMATE	66 distinctions

Every noun in Bora-Miraña has an appropriate GCM. Nouns referring to lower animals and inanimates may in addition have an appropriate SCM. These nouns have available the larger set of SCMs (66 distinctions) and the smaller set of the single inanimate GCM; regardless of the target, the distinctions of the SCM set are consistently neutralized in the GCM set. Thus if a noun is labelled as masculine or feminine (if animate), or according to the appropriate SCMs (if inanimate), then that is sufficient for the syntax to function.

Bora-Miraña is typologically quite unique because both sets of markers are possible for all targets. Compare this with the similar but more expected situation in the eastern Tucanoan language Tatuyo (Gomez-Imbert 2007), spoken by 400 people in Colombia. Like Bora-Miraña, Tatuyo has two sets of markers. One is found on noun phrase modifiers, one is found on the verb. Again, we see that the smaller set involves the neutralization of distinctions made in the larger set. Noun phrase modifiers make a large number of distinctions for inanimate nouns, which all neutralize to *-e* in the verb. Unlike Bora-Miraña, targets differ significantly in Tatuyo. The SCMs used with inanimate nouns only appear in the noun phrase, while the verb neutralizes all of these values into a single inanimate value. In Bora-Miraña, the type of the target is irrelevant. Each target in the language has the same possibilities, and there is a thorough-going option for inanimates to either choose a GCM or an SCM. Therefore, Bora-Miraña and Tatuyo illustrate two different types of neutralization between a larger and a smaller system. The way Tatuyo does it, namely to single out a particular type of

target (or a specific set of targets) where the neutralization occurs, is what we would expect. The kind of relation found in Bora-Miraña is unique so far: here no type of target is singled out, but rather the neutralization is possible with all of them.[16] What we find in Bora-Miraña is an extreme mismatch between assignment and exponence. The language relies on one means of exponence, there is only a single slot for each target; but there is more than one means of assignment (for lower animals and inanimates) and these have to collapse onto each other because they are competing for the same slot.

5 Concurrent systems

It is natural to talk of *the* classification system of a given language. We expect there to be only one. And yet we find languages with concurrent systems of classification, which we discuss here. First we deal with the simplest case, that is, languages which combine two classification systems of the same type, taking as our example Paumarí which has two gender systems (Section 5.1). We then consider a more complex interaction, as found in Mian (Section 5.2). This leads to a discussion of concurrent systems more generally (Section 5.3).

5.1 Two systems of the same type

We find two fully orthogonal gender systems in the Arawan language Paumarí (Chapman and Derbyshire 1991; Aikhenvald 2010), spoken by about 300 people in Peru and Brazil. The first gender system has the values masculine and feminine (just in the singular, in the plural the gender contrast is neutralized). The second system, which works on an absolutive basis, partitions nouns into a *ka-* and a non-*ka-* gender (not interacting with number). In the following example, the controller of agreement is *ojoro* '(female) turtle', which belongs

16 This combination – the same alternative markers being available on all targets, and the many-to-one relation between the two sets – recalls the phenomenon of superclassing (see most recently Meakins and Pensalfini 2016 and references there). The Bora-Miraña system differs in two ways from those languages previously described as showing superclassing. First, in terms of the definition of the phenomena, in superclassing the markers are identical, while Bora-Miraña has one marker which is different (the inanimate singular GCM which corresponds to 66 SCMs). Second, in the examples we have, superclassing has arisen at a later stage of a language which previously had just the larger set of oppositions, while in Bora-Miraña we have the opposite situation, since it appears that the larger set is the more recent.

both to the feminine gender and the *ka-* gender. The verb and the demonstrative agree in gender with the controller (feminine) and the verb also takes the *ka-* marker.

(14) Paumarí
ka-voroni-'a-hi *ida* *ojoro*
KA-fall-ASPECT-THEMATIC.F DEM.F female.turtle(F, KA)
'The turtle fell down.' (Aikhenvald 2010: 241)

There is a semantic core to gender assignment in Paumarí (Aikhenvald 2000: 71) in that human males are masculine, human females are feminine, and higher animals can be of either gender depending on sex. Certain types of nouns are clearly associated with a particular gender value, for example all body parts and the vast majority of artifacts and their parts are feminine, while all celestial bodies and almost all reptiles are masculine. Some types of noun are more evenly distributed across both genders, e.g. birds, plants, and weather phenomena. Assignment to the *ka-* gender is complex: subsets of nouns denoting body parts, plants, fruit, artifacts, and animals; substances which consist of small particles or are thick in texture; and larger flatter objects tend to be in the *ka-* gender (Aikhenvald 2000: 71). No nouns denoting humans are in the *ka-* gender. Thus we have a human versus non-human opposition, with the human (non-*ka-* gender) including some non-humans, but not the contrary.

Taken on its own we would say that the masculine–feminine opposition constitutes a gender system. Equally if we had only the *ka* versus non-*ka* opposition, we might well argue that Paumarí had a gender system in which assignment is partially determined by humanness. It is easy to jump to the conclusion that since there are two distinct oppositions, one must be something else, for example a classifier system. But this does not necessarily follow. It could be that we simply have two gender systems. Or we could have a "composed" gender system (Corbett 2012: 174–180), in which the interaction of the two oppositions gives a gender feature with four values (M-*ka*, M-non-*ka*, F-*ka*, F-non-*ka*).

When evaluating these alternatives, we conclude that in Paumarí we are dealing with two systems. The semantics of the two systems are different. Given the value of any noun in one system we cannot predict its value in the other system. Corbett (2007: 257), citing an earlier version of Aikhenvald (2010), points out that the two systems are independent in that all four combinations are attested (Table 5).

We should add that nouns are not equally distributed over the four possibilities. There is an obvious implication between *ka-* class and gender in that most *ka-* nouns are feminine, while only a few are masculine, e.g. *kasi'i*

Table 5: Paumarí system matrix.

	ka-	non-*ka-*
MASCULINE	*kasi'i* 'alligator'	*makhira* 'man'
FEMININE	*ba'dana* 'lizard'	*arabo* 'land, ground'

'alligator', *vahajari* 'cayman', and *maoba* 'ritual building' (Aikhenvald 2010: 246). Nonetheless the semantics (grammatical meanings) of the two systems are independent of each other, which makes Paumarí an example of a language close to having a canonical arrangement of two systems. What is particularly interesting about Paumarí is that if we had just one of the systems, we would analyse it as a gender system. It seems consistent, therefore, to say that Paumarí has two gender systems. Furthermore, the means of exponence are quite different for the two systems. This means that we have an extreme with both semantics and exponence being different in the two systems of classification.[17]

5.2 Mian

While Paumarí provides an example of a language with two fully orthogonal gender systems, we now turn to the more difficult case of Mian with two systems which are less than fully orthogonal. The Ok language Mian (Trans New Guinea), spoken in Papua New Guinea by 1,700 speakers, has been analysed as combining a four-term gender system with six verbal classifiers (Fedden 2011). The interest and difficulty of Mian is first that these two candidate systems are more similar to each other in their realization than the laconic labels 'gender' and 'classifiers' would suggest, and second that the two systems are far from fully orthogonal. We therefore include Mian as a good illustration of a language where to suggest it had a single system of nominal classification would be false, but to state that it has two, without qualification, would be misleading. We give the data essential to making this point. For fuller detail on the language as a whole the reader is referred to Fedden (2011), and for the systems of nominal classification to Fedden and Corbett (2017a) and especially to Corbett et al. (2017).

The Mian gender system has four values: masculine, feminine, neuter 1, and neuter 2. Examples (15) and (16) illustrate the agreement in gender of the clitic article (with its noun) and the verb (with subject and object).

17 For analysis of the two systems of the related language Kulina, see Dienst (2014: 70–92).

(15) Mian
ē *unáng=o* *wa-têm'-Ø-e=be*
3SG.M woman(F)=ARTICLE.SG.F 3SG.F.OBJ-see.PFV-REALIS-3SG.M.SBJ=DECL
'He sees the woman.'

(16) *ō* *naka=e* *a-têm'-Ø-o=be*
3SG.F man(M)=ARTICLE.SG.M 3SG.M.OBJ-see.PFV-REALIS-3SG.F.SBJ=DECL
'She sees the man.'

Agreement targets are articles and verbs, as illustrated, pronouns and demonstratives. Subject agreement is found on all finite verbs, whereas object agreement is restricted to seven verbs (Fedden 2011: 265–267).

Table 6 sets out the four genders. For exponence we give the forms of the clitic article. (The other targets show different agreement forms but follow exactly the same pattern.)[18] For assignment we give a characterization of the underlying semantics.

Table 6: Mian gender: clitic articles forms and assignment.

	SINGULAR	PLURAL	Assignment
MASCULINE	*=e*	*=i*	Males
FEMININE	*=o*	*=i*	Females
NEUTER 1	*=e*	*=o*	Inanimates
NEUTER 2	*=o*	*=o*	Inanimates: locations, body decoration, weather phenomena, illnesses, abstract nouns, some tools, and weapons

We now turn to the classifiers. These appear only as prefixes on the verb. They operate on an absolutive basis; most of them classify transitive objects and for just one verb ('fall') they classify the intransitive subject. They are restricted to occurring on about 30 verbs of object handling or movement, such as 'give', 'take', 'throw', and 'fall'. The use of the classifier *tob-* 'long object (SG)' for *gíng* 'midrib (of a leaf)' is illustrated in example (17):

18 The agreements show an interesting pattern of syncretism. All Mian genders are non-autonomous values (Zaliznjak 2002[1973]: 69–74, discussed in Corbett 2012: 156–158, 167–170): no gender value has unique agreement forms.

(17) *gíng=e* *tob-tlâa'-n-i=a* [...]
midrib(N1)=ARTICLE.SG.N1 3SG.LONG.OBJ-remove-SS.SEQ-1SG.SBJ=MED [...]
'I remove the midrib (of the tobacco leaf) and then I ...' (Fedden 2011: 541)

As the set of verbs which take object agreement and the set of verbs which take a classifier do not intersect, there is no situation where a single verb would have a classifier and object agreement at the same time.

The classifier system is of enormous interest in typological terms: it has a number opposition (it is orthogonal to number, just as the gender system is, but the interaction works differently). Moreover, the number possibilities are different for animates and inanimates (which have a paucal). For this interesting issue, we refer the reader to Corbett et al. (2017). The essentials, the singular form for exponence (which appears on verbs only) and the assignments are given in Table 7.

Table 7: Mian classifiers: verb prefixes and assignment.

CLASSIFIER	SINGULAR	Assignment
M-CLASSIFIER	*dob-*	males and some inanimates
F-CLASSIFIER	*om-*	females and some inanimates
LONG-CLASSIFIER	*tob-*	long objects
BUNDLE-CLASSIFIER	*gol-*	bundle-like objects
COVERING-CLASSIFIER	*gam-*	covering objects
RESIDUE-CLASSIFIER	*ob-*	residue

What is crucial for our purposes is the relations between the two systems. While the exponence is quite different, the assignments have clear similarities. In fact the degree of orthogonality of the gender system and the classifier system is quite low. Often the value of a noun in one system is predictable from its value in the other. For example, from the perspective of gender, all masculine nouns take the M-CLASSIFIER. From the perspective of the classifiers, all nouns which take the long, covering, or bundle classifier have neuter 1 gender. If we multiply four genders with six classifiers we get 24 theoretically possible combinations of which only nine are attested (Table 8, in which cells filled with examples are the attested combinations, including the counts for each cell based on a dictionary of 894 Mian nouns).

Unlike Paumarí, whose systems of nominal classification show full orthogonality, in Mian the two possible systems are aligned to a considerable extent. In many instances, if nouns were specified with the value of either the gender or the classifier, the other value would follow. We also see that in the three

Table 8: Mian gender and classifiers: orthogonality (Fedden and Corbett 2017a: 16).

	MASCULINE	FEMININE	NEUTER 1	NEUTER 2
M-CLASSIFIER	150, man, boy, boar	–	62, sleeping bag, plate, mosquito net	–
F-CLASSIFIER	–	129, woman, girl, sow	–	205, house, steel axe, money (kina note)
LONG	–	–	142, tobacco, eating implement, bush knife	–
BUNDLE	–	–	2, string bag plastic bag	–
COVERING	–	–	4, blanket, band aid	–
RESIDUE	–	3, tortoise, scorpion	197, cassowary egg, plane, hat	–

cases where this does not hold, the distribution of nouns between genders is rather uneven (M-CLASSIFIER, F-CLASSIFIER) or extremely uneven (RESIDUE): 212 nouns take the M-CLASSIFIER, of these 71% are MASCULINE and 29% are NEUTER 1; 334 nouns take the F-CLASSIFIER, of these 39% are FEMININE and 61% are NEUTER 2; 200 nouns take the RESIDUE CLASSIFIER, of these 2% are FEMININE and 98% are NEUTER 1. On the other hand, we find more than the minimum number of filled cells, which would be six, i.e. the number of values in the larger system (if fewer than six cells were filled, there would be no basis for saying that the classifier system had six values at all). Thus while there are two systems, the situation is quite different from the one we see in Paumarí.[19] For an in-depth account of nominal classification in Mian, including detailed argumentation for and against an analysis of Mian as having two systems, see Corbett et al. (2017).

5.3 Concurrent systems beyond nominal classification

Our concern is nominal classification, but we should situate the analysis of concurrent systems of nominal classification within the wider discussion of concurrent systems. Two papers consider in detail the arguments which can

19 And most recently, Loporcaro (2018: 160–194) argues for a concurrent gender analysis of the Asturian system.

be used to support an analysis as concurrent systems for other linguistic features. Goddard (1982) is concerned with case, and particularly with the type of analysis which used to be common as applied to several languages of Australia (but was rare for other languages with similar systems). This involved split ergativity, with different case systems proposed for nouns and pronouns. Goddard argues convincingly that, in the instances under discussion, there is no justification for proposing concurrent case systems; rather there is an interesting pattern of syncretism. The second paper concerns tense, aspect and mood in Kayardild, for which two previous analyses had been based on concurrent systems. Contrary to this, Round and Corbett (2017) argue that the evidence suggests rather that there is a single system in Kayardild. By examining carefully the arguments for and against concurrent systems in Kayardild, they reach a more satisfying analysis of the data (in this instance as a single system).

Turning now specifically to nominal classification, Fedden and Corbett (2017a) analyse a range of languages with arguably concurrent systems of nominal classification, and give a typology of the possible interactions involved. In a case study based on this typology, Fedden and Corbett (2017b) examine the classifiers of Lao, and argue for a single system here. Finally, Corbett et al. (2017) re-examine the challenging systems of Mian, and provide a means for measuring concurrency, when a sufficient lexicon is available. We cite these papers to emphasize that there are established criteria for proposing single or concurrent systems, and that we do not propose concurrent systems without strong evidence.

6 Extreme concurrency

Given that we can have two concurrent systems of nominal classification, our interest in extremes naturally pushes us to ask whether there can be languages with more than two concurrent systems. Large systems have indeed been claimed to exist in the literature. However, particular claims are not all based on similar analyses. We therefore first discuss how we can evaluate such systems from a common baseline (Section 6.1), and then we analyse three potentially extreme systems (Section 6.2–Section 6.4).

6.1 More than two concurrent systems

The examples of extreme concurrency we examine here are Akatek (Section 6.2), Palikur (Section 6.3), and Carrier (Section 6.4). The numbers of concurrent

systems previously identified are four for Akatek (Zavala 2000), six for Palikur (Aikhenvald and Green 1998), and no less than nine for Carrier (Poser 2005). However, as we will see below, the results of counting can vary dramatically depending on how we count. In a given language, the number of candidate systems which we can identify, based on different semantics or different forms, is not necessarily the same as the number of concurrent systems we want to accept after a full analysis.

First, we need to be aware that candidate systems can be in a many-to-one relation. This is the case when one system makes more fine-grained distinctions where the other system only has a single category. Consider, for instance, a simplified and regularized version of English, in which the pronoun *he* was used for male humans only, *she* for female humans, and *it* for the residue. In the language the relative pronouns were *who* for humans only and *which* for the residue. It is evident that the two candidate systems are not exactly the same, but they are not fully distinct either, in that the relative pronoun system is a simple reduction of the personal pronoun system. We can say that the distinctions in the larger system (the personal pronouns, in our example, with three possibilities) are partially neutralized in the smaller system (the relative pronouns, with two possibilities). In such a situation, the value in the smaller system is always predictable from the value of the larger system. In our terms, this counts as practically the same. For a more detailed justification for treating such many-to-one relations as single systems, see Fedden and Corbett (2017a: 25).

Second, it is not necessarily the case that phonologically different exponents, as markers of values of a classification system on different parts of speech, are evidence of concurrent systems. In other words, it is possible for a language to have a unified system of meaning distinctions, realized through different sets of forms. We use a familiar language to illustrate this point: French has a single gender system, distinguishing masculine and feminine values, but the forms expressing these values depend on the target. In the following example, consider the behaviour of the definite article and the adjective. (A phonemic rendition of French with segmentation is used, in order to avoid confusion induced by the orthography.)

(18) French

le	*garçon*	*est*	*content*
l-ə	gaʁsõ	ɛ	kõtɑ̃
DEF-M.SG	boy(M)[SG]	is	happy[M.SG]

'The boy is happy.'

(19) *la* *femme* *est* *contente*
l-a fam ɛ kõtã-t
DEF-F.SG woman(F)[SG] is happy-F.SG
'The woman is happy.'

The different targets consistently mark a single gender system. The agreement markers on the definite article and the adjective are phonologically different, but what is important here is that they realize the same gender values. There is more variety than our examples imply, since several different final consonants appear on adjectives when feminine. Yet no one would suggest that French had two concurrent gender systems, one realized on the definite article and one realized on the adjective.

We have discussed two situations where there are two candidate systems but where we would not want to speak of concurrent systems. These situations were (i) many-to-one relations between candidate systems and (ii) unified systems of meaning distinctions, which are realized through different sets of forms. While both of them are relevant for languages with simple concurrency, i.e. languages with just two candidate systems, their identification is particularly important in situations of potential extreme concurrency.

6.2 Akatek

Akatek (Mayan, Q'anjob'alan subgroup) is spoken by around 30,000 people in Guatemala and by about 10,000 more in Mexico, the United States and Canada (Zavala 2000: 114). Akatek has been analyzed as having more than two concurrent systems of nominal classification (Zavala 2000, R. Zavala, p.c.), each comprising a set of classifiers. The first candidate is a system of three obligatory suffixal numeral classifiers: for humans the form is *-wan*, for animals it is *-k'on*, and for inanimates *-eb'*. These attach to all numerals (except *jun* 'one') and the interrogative quantifier *jay-* 'how many'. Assignment is strictly semantic. An example is (20).

(20) Akatek
'ey kaa-wan skutzin patron tu'
EXISTENTIAL two-CLF.human daughter boss DIST
'The boss had two daughters.' (Zavala 2000: 118)

The second candidate system is a small system of two plural words which follow a human-nonhuman distinction. We can see quickly that this system subsumes

the first one and there is no reason to say that they constitute two separate classification systems.

The third candidate system consists of 14 noun classifiers. These classify according to sex or animacy (man, woman, animal), familiarity (honorific, familiar) or material/essence (e.g. rock, corn, water, fire, vegetable). An example is (21).

(21) *asi' 'i' eb' naj qetb'i tu'*
go! carry HUMAN.PL CLF.man companion DIST
'Go and bring our companions!' (Zavala 2000: 133)

Finally, there is a fourth candidate system in Akatek consisting of free numeral classifiers. Here the dominant semantic parameter is shape. These classifiers are optional and can only follow a numeral which also has a suffixed numeral classifier from the first candidate system.

So does Akatek really have four concurrent systems: (i) suffixal numeral classifiers, (ii) plural words, (iii) noun classifiers, and (iv) free numeral classifiers? Above we have already said that (i) and (ii) form a single system. Taking a closer look at candidate systems (i/ii) and (iii), we see that the former subsumes the latter. Therefore (i), (ii), and (iii) can be treated as a single system. To a large extent candidate systems (i/ii) and (iii) are in a many-to-one relation: given the classifier in the noun classifier system (candidate iii), the appropriate classifier in the suffixed numeral classifier system (candidate i) can be predicted, which in turn allows the prediction of the correct plural word (candidate ii) So for a noun taking the noun classifier for 'woman' the appropriate suffixal numeral classifier will be *-wan* 'CLF.human', and for a noun taking any of the material noun classifiers the appropriate suffixal numeral classifier will be *-eb'* 'CLF.inanimate'. There are minor exceptions; for example, mushrooms are classified as animals in the noun classifiers but as inanimate in the numeral classifiers. Despite these few mismatches we would treat the suffixal numeral classifiers and the noun classifiers as being in a many-to-one relation and as constituting a single system. But candidate system (iv), i.e. the free numeral classifiers, is not fully subsumed by the other system (i/ii/iii). There is semantic overlap, i.e. in some cases values can be predicted even for candidate system (iv), but the extent of this overlap is not fully clear.

A key point to draw from the discussion of Akatek is that we have to pay attention to how we count concurrent systems. Since one candidate system can subsume another one, we may analyse the number of concurrent systems in the language as lower than the number of candidate systems.

6.3 Palikur

Palikur is a North Arawakan language spoken by over 1,000 people in northern Brazil and French Guiana (Aikhenvald and Green 1998: 429, endnote 1). The first candidate system in Palikur has three values: masculine, feminine and neuter. Nouns denoting humans and some large animals are assigned gender semantically according to sex. Other animals are assigned gender either according to species, e.g. turtles are feminine and insects are masculine, or according to size, whereby large is associated with masculine and small with feminine. Inanimate nouns are assigned to one of masculine, feminine, or neuter, partially determined by shape, consistency and material (Aikhenvald and Green 1998: 436).

Candidate systems 2 to 5 are types of classifier. These systems are based on essentially the same semantics, and in parts use the same forms as well. Candidate 2 consists of numeral classifiers, some of which combine with all numerals from 'one' to 'ten', some only with 'one' to 'two', some only with 'one'. These classifiers are obligatory in the sense that if a numeral can take a classifier it has to. Verbal classifiers (candidate 3) classify the subject of stative verbs of the following semantic types: dimension (e.g. 'be large'), physical property (e.g. 'be smooth'), or colour (e.g. 'be red'). Verbal classifiers (candidate 4) classify the object of transitive verbs of physical action (e.g. 'grab', 'wash', or 'hit'). Unlike any of the other classifiers, verbal classifiers are optional. For locative classifiers (candidate 5), Palikur has a set of forms that function as locative expressions meaning 'on' or 'in', e.g. *-peru* 'on.branchlike', which is the appropriate preposition 'on' for the noun *ah* 'tree'.

Let us look at these four candidate systems which we have identified tentatively above. These candidate systems share a common set of semantic contrasts. We start with the two types of verbal classifier. Table 9 shows that the forms are identical for about half the classes. And for the other half, the forms are often very similar. Sometimes transitive verbs make fewer distinctions than stative verbs, e.g. transitive verbs do not distinguish between 'edge' and 'pointed', whereas stative verbs do. As we have seen for Akatek above, the fact that one candidate system makes more semantic distinctions than another does not necessarily mean that there are two concurrent systems. In fact, we treat cases where one system is subsumed (the one with more distinctions) by another (the one with fewer distinctions) as single systems. We would come to a similar conclusion of 'one system' if we compared the locative classifiers with any of the verbal classifiers, though the degree of formal overlap will be slightly less.

The numeral classifiers are somewhat more involved, but the complication is mainly in the formal realization. They are suffixed to the native numerals 'one'

Table 9: Palikur numeral, verbal (stative and transitive) and locative classifiers (adapted from Aikhenvald and Green 1998)[20].

Semantics	Numeral	Verbal (stative)	Verbal (transitive)	Locative
animate	*-p*	*-pit*	*-pit*	*-pit*
round, square	*-u/-so*	*-pit*	*-pit*	*-pit*
irregular shape	*-a/-sa*	*-pit*	*-pit*	*-pit*
side	*-a/-sa*	*-muh*	*-muh*	*-pit*
vertical objects	*-t/-ta-*	*-min*	*-min*	*-min*
rigid, thin	*-t/-ta-*	*-ah*	*-min*	*-min*
flat	*-k/-ka-/-bu*	*-boha*	*-bo*	*-madka*
concave	*-mku/-muk*	*-apa*	*-ap*	*-madka*[21]
edge	*-mku/-muk*	*-kiya*	*-kig*	*-kigbi*
pointed	*-mku/-muk*	*-kisa*	*-kig*	*-kigsa*
linear	*-tra/-tahr-/-bu*	*-buka*	*-buk*	*-buhku(mna)*
road, river	*-tra/-tahr-/-bu*	*-buka*	*-buk*	*-vigku*
inside part of	*-iku/-rik*	*-eku*	*-ik*	*-iku*
tree, plant, trunk	*-kti/-kat*	*-kat*	*-min*	*-pew*
tree, branchlike	*-kti/-kat*	*-pewa*	*-peru*	*-peru*
water	–	*-pit*	*-pit*	*-hakwa*

and 'three' to 'ten', and infixed into 'two', using slightly different forms. Compared to the verbal and locative classifiers, the forms are different, but the semantics are (almost) the same. Importantly, in such a situation, we should not conclude that we are dealing with two concurrent systems. Above we used French to illustrate this point.

For Palikur, if we compare the numeral classifiers pairwise with the other classifiers, we find that each relation is one of same semantics but different forms. So we conclude that the classifiers in Table 9 constitute one system.

Candidate system 6 is a system of five possessive classifiers, whose forms are given in Table 10. An example is *pi-mana uwas* [2SG-CL.food orange] 'your orange' (Aikhenvald and Green 1998: 460).

20 For numeral classifiers, the first form only appears on 'one', the second only appears on 'two' and the third only appears on numerals above 'two'. If no third form is given, that classifier is not available for the numerals 'three' and above. If only one form is given, the classifier only appears with 'one'.

21 Table 8 in Aikhenvald and Green (1998: 462) has *-madka* here, but Figure 5 on locative classifiers on page 457 has *-apa* instead for this value. Aikhenvald (2000) gives only *-madka* for the concave locative classifier.

Table 10: Palikur possessive classifiers (Aikhenvald and Green 1998: 460).

Classifier	Meaning
-pig	'pet' (for domestic animals)
-mana	'food' (for fruit and vegetables)
-mutra	'plant'
-win	'catch, animal caught to be eaten'
-kamkayh	'child'

A comparison of the possessive classifiers (candidate 6) with the classifiers in Table 9 (candidates 2 to 5) and with the gender agreements (candidate 1) shows that the formal realizations are different. Since different formal realizations are not a sufficient condition to assume concurrent systems we need to turn to the semantics. There seems to be at least some overlap, though the exact extent is not clear from the literature on Palikur. As far as Palikur gender is concerned, any pairwise comparison of gender with any of the classifier sets might yield fairly independent systems because notional inanimates can be masculine, feminine or neuter, which means that any gender value can combine with the classifiers in Table 9 and possibly the possessive classifiers in Table 10, but again the degree of semantic overlap is not fully clear. Thus we conclude that Palikur has maximally three concurrent classification systems; depending on the degree of semantic overlap there might be fewer. Again we see that it is essential that we calibrate our count of concurrent systems.

6.4 Carrier

The Athabaskan language Carrier, which has 11,000 speakers in Canada, is famous for its richness in nominal classification systems. Poser (2005: 162–163) identifies no fewer than nine subsystems in Carrier, which he calls generally incommensurable. This would clearly qualify as extreme concurrency of systems. As the situation in Carrier is complicated we highlight aspects of it and for each we will show that we have to argue carefully for the relations between candidate systems.

The first candidate system is a set of three absolutive cross-referencing verb prefixes with shape semantics: *d-* 'sticklike object', *n-* 'round object', and x^w- 'spatially extended object', which can be found throughout the Athabaskan family (Davidson et al. 1963). There is an unmarked category, which is called *generic* in the Athabaskan tradition. Humans, for example, belong to the generic class. Assignment is based on semantics, but there is opacity, for example

sounds (e.g. message, speech, song, or word) take the *d-* form. Example (22) illustrates the use of *x*^{w}- for *yoh* 'house', which is classified as a spatially extended object and which appears as both the S argument of the first verb (constituting a relative clause) and the O argument of the second.

(22) Carrier (Poser 2005: 145)[22]
Yoh *x*^{w}*ʌdiz̲k'ʌnʌn* *x*^{w}*ʌnɨɬʔen.*
house which.x^{w}.is.burning he.is.looking.at.x^{w}
'He is looking at the burning house.'

The second candidate system consists of cross-referencing prefixes on postpositions and possessed nouns. There is the familiar prefix *x*^{w}- 'spatially extended object', whereas the other three categories are covered by a single form (*b-*). Although the forms are slightly different we still see that the first candidate system is in a many-to-one relation with the second. We would therefore say that they are a single system.

The third candidate system consists of a set of 11 classificatory verbs; as in many Athabaskan languages, these are used to describe the handling, location, and motion of objects of different types.

Table 11 is the system matrix for Carrier which shows the semantic overlap between the absolutive prefixes (top) and the classificatory verbs (down the side, specifying the different types of object involved in the handling). "+" indicates that there are nouns which take the given combination.

The formal realizations of the candidate systems are entirely different. While none of the systems is completely predictable from the other (in other words, neither of the systems are fully subsumed by the other) there is considerable semantic overlap. Given the classificatory verbs (left column in Table 11) we can predict the correct absolutive prefixes (top row in Table 11) in most cases. This only fails for the classificatory verbs for 'non-plural default' and 'long rigid', which can combine with more than one absolutive prefix.

We do not have the space to go through the remaining candidate systems which Poser identifies. Our point is rather that if we carried on comparing candidate systems with each other, as we did for Mian (Section 5.2), Akatek (Section 6.2) and Palikur (Section 6.3) we would end up with four systems of nominal classification, which cannot be further reduced by subsuming one within another. But this is not to say that Carrier has four completely independent systems, the way that Paumarí has two completely independent systems (Section 5.1). In each pairwise comparison between the four remaining systems of Carrier there is always semantic

22 Carrier forms are not segmented in Poser (2005).

Table 11: Carrier system matrix (adapted from Poser 2005: 151)[23].

	Ø- (generic)	*d-*	*n-*	x^w-
non-plural object	+	+	+	+
long rigid	+	+	–	–
body	+	–	–	–
contents of open container	+	–	–	–
2D flexible	+	–	–	–
mushy	+	–	–	–
fluid	+	–	–	–
hay	–	+	–	–
fluff	–	–	+	–

overlap (sometimes substantial). So Carrier is a very clear illustration of the importance of being explicit about how we determine the number of systems.

7 Conclusion

We have established the extremes of nominal classification systems, from the surprisingly simple to the surprisingly complex. With this cline in mind, we looked at the essential components of any nominal classification system, i.e. assignment and exponence. We then considered extreme configurations of assignment and exponence in individual languages. The range of variation that we established is considerable: given that a language may have the extremely simple gender assignment system of Bagvalal, the extremely complex system of German is the more surprising. And given Rosch's point cited earlier, that categorization systems should "provide maximum information with the least cognitive effort" (1978: 190), it is significant that we find both extremes, Bagvalal and German, as functioning systems. Thus we claim that our analysis is important for research into categorization, and for cognitive linguistics generally. In particular, the complexity of some of the systems investigated is all the more remarkable when we see just how simple nominal classification systems can be.

There is, however, a whole further type of extreme. While in the simple case, a language has a single classification system, we have set out several more

23 We have omitted the two classes which are concerned with number (plural default, which contains the plurals of the default category) and countability (effectively uncountable, which contains mass nouns like 'sugar' and 'berries').

extreme cases where we find languages with two or even more concurrent systems. The study of these concurrent systems is potentially highly significant. They offer additional experimental opportunities, since they allow many troublesome variables to be held constant. Since the systems coexist in a single language, all the rest of the language, from its phonetics to its sociolinguistic situation, can be held constant, while we work on and compare just their concurrent systems of nominal classification.

For typology, the major contribution is in further developing this recent research into concurrent systems. Within nominal classification, we take forward the analyses of Seifart (2005), Fedden (2011), and Fedden and Corbett (2017a).

Most importantly, our research brings these two strands – cognition and typology – together. This can be seen as a specific type of link, with typology providing the essential parameters for experiments. Until quite recently, the interests of cognitive science and typology were rather distinct. But the situation has changed dramatically: experiments that could once be carried out only in laboratories, using large and expensive equipment, can increasingly be carried out in the field, even in remote locations, using portable and much less costly equipment (Harris and Samuel 2011; Harris 2017: 91–95). We have to learn to negotiate the trade-off between data obtained in conditions which are not as perfectly controlled as would be ideal, with the possibility of investigating a much wider range of phenomena in a greatly expanded set of languages. Specifically in respect to nominal classification, effects such as the gender congruency effect (Schriefers 1993; discussed in Schiller 2014) can be investigated in languages of different genetic and areal distribution (see, for instance, Tsegaye et al. 2014 on Konso). We thus concentrate on a particular connection between typology and cognition. For broader discussion of typology and cognitive linguistics, see Croft (2016).

The exciting new possibilities bring with them two distinct challenges. The first is the need for great care over experimental design. If there is to be a real strengthening of the connection between linguistic data and experimental psychology, we need to be confident about what the measurements mean. The second is that those same measurements must be consistent across languages, and it is here that our canonical approach to typology has its place. We have been constructing a typology of nominal classification that others can work with. At the most basic level, we need to be assured that when we claim that a particular system of nominal classification draws more distinctions than another (is more extreme than another), we have the typological means to prove it.[24]

24 Linguists are sometimes unwilling to give due care to measurement and consistency. While the cost of linguistic errors of measurement may be lower, we can all take a lesson from the

By analysing in detail the extremes of nominal classification, the simple as well as the complex, we show the richness of the 'natural laboratory' which the world's languages offer. It would be foolish to ignore it, and also foolish to devote our experimental work to one limited corner of this extensive space.

Abbreviations

1	first person
1D	one-dimensional
3	third person
I	gender I
II	gender II
III	gender III
IV	gender IV
ABS	absolutive
ATTR	attributive
DAT	dative
CLF	classifier
COP	copula
DECL	declarative
DEF	definite
DEM	demonstrative
DIST	distal
ERG	ergative
F	feminine
GCM	general class marker
INAN	inanimate
INCL	inclusive
M	masculine
N	neuter
OBJ	object
PFV	perfective
PL	plural
POSS	possessive
SBJ	subject
SCM	special class marker
SG	singular

NASA Mars Climate Orbiter disaster on Mars, September 1999, when the failure to be consistent over units of measurement led to a substantial error, and the loss of the orbiter.

Acknowledgements: The support of the AHRC (grant AH/N006887/1: Lexical splits: a novel perspective on the structure of words) is gratefully acknowledged. We are grateful to four anonymous referees for *Cognitive Linguistics* for their suggestions. We thank Penny Everson and Lisa Mack for their help in preparing the manuscript. We are joint authors and the order of names is not significant.

References

Acuña-Fariña, Carlos. 2016. (A few) psycholinguistic properties of the NP. *Functions of Language* 23. 120–141.

Aikhenvald, Alexandra Y. 2000. *Classifiers: A typology of noun classification devices*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2010. Gender, noun class and language obsolescence: The case of Paumarí. In Eithne B. Carlin & Simon van de Kerke (eds.), *Linguistics and archeology in the Americas: The historization of language and society*, 235–252. Leiden: Brill.

Aikhenvald, Alexandra Y. & Diana Green. 1998. Palikur and the typology of classifiers. *Anthropological Linguistics* 40. 429–480.

Asher, R. E. 1989 [1985]. *Tamil*. London: Croom Helm [Routledge].

Audring, Jenny. 2014. Gender as a complex feature. In Tania Paciaroni, Michele Loporcaro & Anna M. Thornton (eds.), *Exploring grammatical gender [Special issue] Language Sciences*, vol. 43. 5–17. http://dx.doi.org/10.1016/j.langsci.2013.10.003 (last accessed 28 August 2018).

Audring, Jenny. 2017. Calibrating complexity: How complex is a gender system? *Language Sciences* 60. 53–68.

Austin, Peter K. 2013 [1981]. *A grammar of Diyari, South Australia* [Cambridge Studies in Linguistics 32], 2nd edn., [Cambridge Studies in Linguistics 32] London: SOAS, University of London [Cambridge & New York: Cambridge University Press].

Baechler, Raffaela & Guido Seiler (eds.). 2016 *Complexity, isolation, and variation*. Berlin: De Gruyter.

Beller, Sieghard, Karen F. Brattebrø, Kristina O. Lavik, Rakel D. Reigstad & Andrea Bender. 2015. Culture or language: What drives effects of grammatical gender? *Cognitive Linguistics* 26. 331–359.

Bender, Andrea, Sieghard Beller & Karl Christoph Klauer. 2016. Lady Liberty and Godfather Death as candidates for linguistic relativity? Scrutinizing the gender congruency effect on personified allegories with explicit and implicit measures. *The Quarterly Journal of Experimental Psychology* 69. 48–64.

Binanzer, Anja. 2017. *Genus – Kongruenz und Klassifikation: Evidenzen aus dem Zweitspracherwerb des Deutschen*. Berlin: De Gruyter Mouton.

Bond, Oliver. 2013. A base for canonical negation. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 20–47. Oxford: Oxford University Press.

Bond, Oliver, Greville G. Corbett & Marina Chumakina. 2016. Introduction. In Oliver Bond, Greville G. Corbett, Marina Chumakina & Dunstan Brown (eds.), *Archi: Complexities of*

agreement in cross-theoretical perspective (Oxford Studies of Endangered Languages), 1–16. Oxford: Oxford University Press.
Brown, Dunstan & Marina Chumakina. 2013. What there is and what there might be: An introduction to Canonical Typology. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 1–19. Oxford: Oxford University Press.
Burrow, T. & S. Bhattacharya. 1953. *The Parji language: A Dravidian language of Bastar*. Hertford: Austin.
Chapman, Shirley & Desmond C. Derbyshire. 1991. Paumarí. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 3. 161–352. Berlin: Mouton de Gruyter.
Chumakina, Marina & Oliver Bond. 2016. Competing controllers and agreement potential. In Oliver Bond, Greville G. Corbett, Marina Chumakina & Dunstan Brown (eds.), *Archi: Complexities of agreement in cross-theoretical perspective* (Oxford Studies of Endangered Languages), 77–117. Oxford: Oxford University Press.
Chumakina, Marina & Greville G. Corbett. 2008. Archi: The challenge of an extreme agreement system. In A. V. Arxipov, L. M. Zaxarov, A. A. Kibrik, A. E. Kibrik, I. M. Kobozeva, O. F. Krivnova, E. A. Ljutikova & O. V. Fëdorova (eds.), *Fonetika i nefonetika: K 70-letiju Sandro V. Kodzasova* [Phonetics and non-phonetics: For the 70th birthday of Sandro V. Kodzasov], 184–194. Moscow: Jazyki slavjanskix kul´tur.
Chumakina, Marina & Greville G. Corbett. 2015. Gender-number marking in Archi: Small is complex. In Matthew Baerman, Dunstan Brown & Greville G. Corbett (eds.), *Understanding and measuring morphological complexity*, 93–116. Oxford: Oxford University Press.
Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
Corbett, Greville G. 2007. Gender and noun classes. In Timothy Shopen (ed.), *Language typology and syntactic description: III: Grammatical categories and the lexicon*, 2nd edn., 241–279. Cambridge: Cambridge University Press.
Corbett, Greville G. 2012. *Features*. Cambridge: Cambridge University Press.
Corbett, Greville G. 2014. Gender typology. In Greville G. Corbett (ed.), *The expression of gender* (The Expression of Cognitive Categories [ECC] 6), 87–130. Berlin: De Gruyter.
Corbett, Greville G. 2015. Morphosyntactic complexity: A typology of lexical splits. *Language* 91. 145–193.
Corbett, Greville G. & Sebastian Fedden. 2016. Canonical gender. *Journal of Linguistics* 52. 495–531.
Corbett, Greville G., Sebastian Fedden & Raphael A. Finkel. 2017. Single versus concurrent feature systems: Nominal classification in Mian. *Linguistic Typology* 21. 209–260.
Corbett, Greville G. & Alfred D. Mtenje. 1987. Gender agreement in Chichewa. *Studies in African Linguistics* 18. 1–38.
Croft, William. 2016. Typology and the future of Cognitive Linguistics. *Cognitive Linguistics* 27. 587–602.
Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins.
Davidson, William, William Elford & Harry Hoijer. 1963. Athapaskan classificatory verbs. In Harry Hoijer (ed.), *Studies in the Athapaskan languages*, 30–41. Berkeley: The University of California Press.

Demuth, Katherine A. 1988. Noun classes and agreement in Sesotho acquisition. In Michael Barlow & Charles Ferguson (eds.), *Agreement in natural language: Approaches, theories, descriptions*, 305–321. Stanford: CSLI.

Demuth, Katherine. 2003. The acquisition of Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 209–222. London: Routledge.

Di Garbo, Francesca & Bernhard Wälchli (eds.), Forthcoming. *Grammatical gender and linguistic complexity*.

Dienst, Stefan. 2014. *A grammar of Kulina* (Mouton Grammar Library 66). Berlin: De Gruyter Mouton.

Dixon, R. M. W. 1972. *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1982. Nominal classification. In R. M. W. Dixon (ed.), *Where have all the adjectives gone? and other essays in semantics and syntax*, 157–233. Berlin: De Gruyter.

Dixon, R. M. W. 1986. Noun classes and noun classification in typological perspective. In Colette Craig (ed.), *Noun classes and categorization: Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983*, 105–112. Amsterdam: John Benjamins.

Enger, Hans-Olav. 2009. The role of core and non-core semantic rules in gender assignment. *Lingua* 119. 1281–1299.

Enger, Hans-Olav. 2011. Gender and contact: A natural morphology perspective on Scandinavian examples. In Peter Siemund (ed.), *Linguistic universals and language variation*, 171–203. Berlin: De Gruyter Mouton.

Evans, Nicholas, Henrik Bergqvist & Lila San Roque. 2018. The grammar of engagement II: Typology and diachrony. *Language and Cognition* 10. 141–170.

Evans, Nicholas, Dunstan Brown & Greville G. Corbett. 2002. The semantics of gender in Mayali: Partially parallel systems and formal implementation. *Language* 78. 111–155.

Fedden, Sebastian. 2011. *A grammar of Mian* (Mouton Grammar Library 55). Berlin: De Gruyter Mouton.

Fedden, Sebastian & Greville G. Corbett. 2017a. Gender and classifiers in concurrent systems: Refining the typology of nominal classification. *Glossa: a Journal of General Linguistics* 2(1). 34. http://doi.org/10.5334/gjgl.177 (last accessed 28 August 2018).

Fedden, Sebastian & Greville G. Corbett. 2017b. Understanding intra-system dependencies: Classifiers in Lao. In N. J. Enfield (ed.), *Dependencies in language*, 173–181. Berlin: Language Science Press.

Forker, Diana. 2016. Conceptualization in current approaches of language typology. *Acta Linguistica Hafniensia* 48. 70–84.

Franjieh, Michael J. 2012. *Possessive classifiers in North Ambrym, a language of Vanuatu: Explorations in semantic classification*. London: University of London, School of Oriental and African Studies (SOAS) thesis.

Franjieh, Michael J. 2018. North Ambrym possessive classifiers from the perspective of canonical gender. In Sebastian Fedden, Jenny Audring & Greville G. Corbett (eds.), *Non-canonical gender systems*, 36–67. Oxford: Oxford University Press.

Gagliardi, Annie, Naomi H. Feldman & Jeffrey Lidz. 2017. Modeling statistical insensitivity: Sources of suboptimal behavior. *Cognitive Science* 41. 188–217.

Gagliardi, Annie & Jeffrey Lidz. 2014. Statistical insensitivity in the acquisition of Tsez noun classes. *Language* 90(1). 58–89.

Goddard, Cliff. 1982. Case systems and case marking in Australian languages: A new interpretation. *Australian Journal of Linguistics* 2. 167–196.

Gomez-Imbert, Elsa. 1982. *De la forme et du sens dans la classification nominale en Tatuyo (langue Tukano orientale d'Amazonie colombienne)* (Travaux et Documents Microédités 19). Paris: ORSTOM.

Gomez-Imbert, Elsa. 2007. Tukanoan nominal classification: The Tatuyo system. In Leo Wetzels (ed.), *Language endangerment and endangered languages: Linguistic and anthropological studies with special emphasis on the languages and cultures of the Andean-Amazonian border area*, 401–428. Leiden: Research School of Asian, African and Amerindian Studies (CNWS), Universiteit Leiden.

Grinevald, Colette & Frank Seifart. 2004. Noun classes in African and Amazonian languages. *Linguistic Typology* 8. 243–285.

Harris, Alice C. 2017. *Multiple exponence*. Oxford: Oxford University Press.

Harris, Alice C. & Arthur G. Samuel. 2011. Perception of exuberant exponence in Batsbi: Functional or incidental? *Language* 87. 447–469.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Jagersma, Abraham H. 2010. A descriptive grammar of Sumerian. Leiden: Leiden University dissertation.

Kibrik, A. E. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka: II: Taksonomičeskaja grammatika* [A structural description of Archi II: Taxonomic grammar]. (Publikacii otdelenija strukturnoj i prikladnoj lingvistiki [Publications of the Department of Structural and Applied Linguistics] 12). Moscow: Izdatel´stvo Moskovskogo universiteta [Moscow University Press].

Kibrik, A. E., S. V. Kodzasov, I. P. Olovjannikova & D. S. Samedov. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka: I: Leksika, fonetika* [A structural description of Archi: I: Lexis, phonetics] (Publikacii otdelenija strukturnoj i prikladnoj lingvistiki [Publications of the Department of Structural and Applied Linguistics] 11). Moscow: Izdatel´stvo Moskovskogo universiteta [Moscow University Press].

Kibrik, Aleksandr E., K. I. Kazenin, E. A. Ljutikova & S. G. Tatevosov (eds.). 2001. *Bagvalinskij jazyk: Grammatika: Teksty: Slovari* [Bagvalal: Grammar, texts, dictionaries]. Moscow: Nasledie.

Killian, Don. 2015. Topics in Uduk phonology and morphosyntax. Helsinki: University of Helsinki dissertation.

Koenig, Jean-Pierre. 1999. *Lexical relations* (Stanford Monographs in Linguistics). Stanford: CSLI.

Koestler, Arthur. 1983. *Janus. A summing up*. London: Pan Picador.

Köpcke, Klaus-Michael. 1982. *Untersuchungen zum Genussystem der deutschen Gegenwartssprache*. Tübingen: Niemeyer.

Köpcke, Klaus-Michael & David A. Zubin. 1984. Sechs Prinzipien für die Genuszuweisung im Deutschen. Ein Beitrag zur natürlichen Klassifikation. *Linguistische Berichte* 93. 26–50.

Kraaikamp, Margot. 2017. The diachrony of semantic gender agreement: Findings from Middle Dutch. *Journal of Germanic Linguistics* 29(3). 259–297.

Kwon, Nahyun & Erich R. Round. 2015. Phonaesthemes in morphological theory. *Morphology* 25. 1–27.

Lakoff, George. 1987. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lichtenberk, Frank. 1983. Relational classifiers. *Lingua* 60. 147–176.

Lohndal, Terje & Marit Westergaard. 2016. Grammatical gender in American Norwegian heritage language: Stability or attrition? *Frontiers in Psychology* 7. 344. https://doi.org/10.3389/fpsyg.2016.00344 (last accessed 28 August 2018)

Loporcaro, Michele. 2018. *Gender from Latin to Romance: History, geography, typology*. Oxford: Oxford University Press.

Meakins, Felicity & Rob Pensalfini. 2016. Gender bender: Disagreement in Jingulu noun class marking. In Felicity Meakins & Carmel O Shannessy (eds.), *Loss and renewal: Australian languages since colonisation*, 425–450. Berlin: De Gruyter Mouton.

Mel´čuk, Igor A. 2013. The notion of inflection and the expression of nominal gender in Spanish. *Studies in Language* 37. 736–763.

Michael, Lev. 2014. The Nanti reality status system: Implications for the typological validity of the realis/irrealis contrast. *Linguistic Typology* 18(2). 251–288.

Nesset, Tore. 2006. Gender meets the usage-based model: Four principles of rule interaction in gender assignment. *Lingua* 116. 1369–1393.

Nikolaeva, Irina. 2013. Unpacking finiteness. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 99–122. Oxford: Oxford University Press.

Passer, Matthias B. 2016. The typology and diachrony of nominal classification. Amsterdam: University of Amsterdam dissertation.

Pawley, Andrew K. 2002. Using *He* and *She* for inanimate referents in English: Questions of grammar and world view. In N. J. Enfield (ed.), *Ethnosyntax: Explorations in grammar and culture*, 110–137. Oxford: Oxford University Press.

Plaster, Keith & Maria Polinsky. 2010. Features in categorization, or a new look at an old problem. In Anna Kibort & Greville G. Corbett (eds.), *Features: Perspectives on a key notion in linguistics*, 109–142. Oxford: Oxford University Press.

Poser, William J. 2005. Noun classification in Carrier. *Anthropological Linguistics* 47(2). 143–168.

Rădulescu, Roxana & Katrien Beuls. 2016. Modelling pronominal gender agreement in Dutch: From a syntactic to a semantic strategy. *Belgian Journal of Linguistics* 30. 219–250.

Reid, Nicholas. 1997. Class and classifier in Ngan'gityemerri. In Mark Harvey & Nicholas Reid (eds.), *Nominal classification in Aboriginal Australia*, 165–228. Amsterdam: John Benjamins.

Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch & Barbara B. Lloyd (eds.), *Cognition and categorizaion*, 27–48. Hillsdale, NJ: Lawrence Erlbaum.

Round, Erich R. & Greville G. Corbett. 2017. The theory of feature systems: One feature versus two for Kayardild tense-aspect-mood. *Morphology* 27(1). 21–75. https://doi.org/10.1007/s11525-016-9294-3 (last accessed 28 August 2018).

Schiller, Niels O. 2014. Psycholinguistic approaches to the investigation of grammatical gender. In Greville G. Corbett (ed.), *The expression of gender* (The Expression of Cognitive Categories [ECC] 6), 161–189. Berlin: De Gruyter.

Schriefers, Herbert. 1993. Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19. 841–850.

Seifart, Frank. 2005. *The structure and use of shape-based noun classes in Miraña (North West Amazon)*. Nijmegen: Radboud University dissertation.

Seifart, Frank. 2009. Multidimensional typology and Miraña class markers. In Patience Ebbs & Alexandre Arkhipov (eds.), *New challenges in typology: Transcending the borders and refining the distinctions*, 365–385. Berlin: Mouton de Gruyter.

Seifart, Frank. 2018. The semantic reduction of the noun universe and the diachrony of nominal classification. In William B. McGregor & Søren Wichmann (eds.), *The diachrony of classification systems*, 9–32. Amsterdam: John Benjamins.

Senft, Gunter. 1986. *Kilivila: The language of the Trobriand Islanders* (Mouton Grammar Library 3). Berlin: De Gruyter Mouton.
Senft, Gunter. 1993. A grammaticalization hypothesis on the origin of Kilivila classificatory particles. *Sprachtypologie und Universalienforschung* 46. 100–112.
Senft, Gunter. 1996. *Classificatory particles in Kilivila* (Oxford Studies in Anthropological Linguistics). New York: Oxford University Press.
Senft, Gunter. 2007. Nominal classification. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 676–725. Oxford: Oxford University Press.
Siemund, Peter. 2008. *Pronominal gender in English: A study of English varieties from a cross-linguistic perspective*. London: Routledge.
Speed, Laura J., Jidong Chen, Falk Huettig & Asifa Majid. 2016. Do classifier categories affect or reflect object concepts? In A. Papafragou, D. Grodner, D. Mirman & J. Trueswell (eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society (CogSci 2016)*, 2267–2272. Austin, TX: Cognitive Science Society.
Stump, Gregory T. 2016. *Inflectional paradigms: Content and form at the syntax-morphology interface*. Cambridge: Cambridge University Press.
Taylor, John R. 1989. *Linguistic categorization: Prototypes in linguistic theory*. Oxford: Clarendon.
Thornton, Anna M. 2009. Constraining gender assignment rules. *Language Sciences* 31. 14–32.
Tsegaye, Mulugeta T., Maarten Mous & Niels O. Schiller. 2014. Plural as value of Cushitic gender: Evidence from gender congruency effect experiments in Konso (Cushitic). In Greville G. Corbett (ed.), *The expression of gender* (The Expression of Cognitive Categories [ECC] 6), 191–214. Berlin: De Gruyter.
Zaliznjak, Andrej A. 1964. K voprosu o grammatičeskix kategorijax roda i oduševlennosti v sovremennom russkom jazyke [The grammatical categories of gender and animacy in modern Russian]. *Voprosy Jazykoznanija [Issues in Linguistics]* 4. 25–40.
Zaliznjak, Andrej A. 1973. O ponimanii termina ‘padež’ v lingvističeskix opisanijax [Interpreting the term ‘case’ in linguistic descriptions]. In Andrej A. Zaliznjak (ed.), *Problemy grammatičeskogo modelirovanija* [Problems of grammatical modelling], 53–87. Moscow: Nauka. [Problems of grammatical modelling].
Zavala, Roberto. 2000. Multiple classifier systems in Akatek (Mayan). In Gunter Senft (ed.), *Systems of nominal classification*, 114–146. Cambridge: Cambridge University Press.