Beyond the classifier/gender dichotomy

The role of flexibility in a more integrated typology of nominal classification

RUTH SINGER

5.1 The classifier/gender dichotomy and typologies of nominal classification

While Corbett’s notion of ‘canonical agreement’ (Corbett 2006b) has already been applied in studies of nominal classification, the canonical gender approach extends this to nominal classification specifically, describing a fictional gender system that can then be used as a point of comparison to explore the full diversity of the world’s nominal classification systems (Corbett and Fedden 2016; see also Corbett and Fedden this volume). This chapter will discuss the potential that ‘canonical gender’ has to generate new ideas about the typology of nominal classification. Canonical gender will be discussed both in relation to earlier approaches to the typology of gender and the analysis of gender in Mawng, an Australian language.

Nominal classification systems are extremely widespread. Around 40% to 70% of all languages have some kind of recognized nominal classification system.1 In addition, it is now known that many languages have more than one kind of nominal classification system (Fedden 2011; Singer 2016: chapter 7). So the total number of nominal classification systems found around the world may be only a little lower

---

1 This figure is based on an estimate that 40% of the world’s languages have grammatical gender and an estimate that a further 30% may have classifiers. Corbett (2013a) found that 112 out of the 256 languages in his WALS survey had grammatical gender, or 44%. Dahl (2004) agrees, estimating that 24%–40% of the world’s languages have gender, depending on how you define it. Since classifiers are overwhelmingly found in less well-studied languages, the number of languages with classifiers could easily be as high as 30%.
than the total number of languages. Typologies of nominal classification from Dixon (1968) to Grinevald and Seifart (2004) treat gender and classifiers as two separate phenomena, or as distant points on a linear continuum. Effectively, these typologies such as Dixon (1982) separate better-known European-like gender systems from less well-known classifier systems such as East Asian numeral classifiers, Oceanic possessive classifiers, the ‘noun classifiers’ of Meso-American languages, and the quite different ‘noun classifiers’ of Australian languages.

Kilarski’s (2013) recent history of work on nominal classification demonstrates how much more we know about European gender systems than any other kinds of nominal classification. There has been a concerted effort by linguists to analyse classifier systems since at least the 1970s but these new analyses have not yet influenced theoretical debates to any significant extent. We now know much more about diverse nominal classification systems but theoretical work is lagging behind.

Corbett and Fedden (2016) join the growing call to do away with the binary opposition between gender and classifier2 systems (Kilarski 2013; Singer 2016). The weight of evidence from languages with diverse nominal classification systems has clearly shifted the balance. Maintaining the distinction between genders and classifiers has become untenable. The clustering of features claimed in earlier typologies to distinguish gender and classifiers has not held up in the face of the analyses of new systems in the past few decades. In addition, the distinct sets of properties proposed to distinguish genders from classifiers, do not correlate to the extent we might expect. For example, locus of marking and agreement do not predict the flexibility of gender agreement in Mawng (Singer 2016). Similarly, semantic transparency of assignment does not predict whether flexibility or cross-classification is possible (see Franjieh this volume; Singer 2016: chapter 7). Cross-classification is used to describe a situation in which the same referent or noun can be categorized as part of different classes (referred to as ‘recategorisation’ in Corbett and Fedden 2016). The prevalence of cross-classification as defined above is referred to as the ‘flexibility’ of the system.

The emphasis on the flexibility of nominal classification is the major advance of the canonical gender approach which enables it to move the field towards a more integrated approach to the typology of nominal classification—that is, one that is not hobbled by the gender/classifier dichotomy. While the traditional departure point of typologies is the division between gender systems and classifier systems, canonical gender makes flexibility the foundation stone for the typology. Flexibility turns out to be a much more subtle but also more powerful heuristic, as the discussion of Mawng gender in this chapter shows.

2 I use ‘classifier’ here in a way it is often used—to refer to any system of nominal classification apart from gender. The use of this term already hints at some of the issues in the typology of nominal classification: that gender systems are always central to our typologies. Ideally, we would use the term ‘classifier’ more broadly to refer to any nominal classification system, the term gender to refer to a gender system, and the more specific term non-gender-like classifier to refer to any system apart from gender.
Mawng gender is a system of nominal classification that has been argued to challenge the binary opposition between genders and classifiers (Singer 2012, 2016). Although Mawng’s system of nominal classification meets the formal criteria for a gender system, it exhibits unexpected flexibility. The creative ways that the system is used has parallels in the uses of noun classifiers of nearby Australian languages. Indeed, the way the system is used is more similar to noun classifiers in Australian languages than how other gender systems are used (Singer 2016). Using the example of the Mawng gender system, I argue that in order to get closer to an integrated typology of nominal classification, we need to reduce our emphasis on the role of nouns and take seriously other phenomena that drive classification, such as semantic fields and contextual properties of referents. These are often more obviously relevant to nominal classification in languages with classifier systems but play a role in all systems of nominal classification. I argue that since nouns are fairly infrequent in naturalistic discourse in many of the world’s languages, it is not realistic to model nouns as the prime motivators of nominal classification.

Making a clean break from long-held foundations like the gender/classifier dichotomy is never easy. Historically, gender in particular has been analysed as a property of nouns and nominal classification understood primarily as a sort of static system whereby nouns are classified into discrete categories. Agreement is then construed as derived from this ‘basic’ classification. However, rather than agreement being the expression of some pre-existing classification system in fact the system only comes into being and is maintained through its use (cf. Bybee 2010). In this chapter I look at gender in Mawng from two quite different perspectives, which I dub the ‘classical perspective’ and the ‘language-in-use’ perspective. The aim of this is to illustrate further what is missed by a classical approach to nominal classification. The classical perspective is the only perspective we can gain for many lesser-known languages, for which we simply do not have enough data from naturalistic discourse to take the analysis any further.

In sections 5.3–5.6 I demonstrate that the classical view combines well with a noun-centred approach to the use of gender in Mawng. However, as mentioned earlier, the breakthrough of the canonical gender approach is making flexibility of use central to the typology. It is only possible to understand how and in what ways a system is flexible through engaging more broadly with how speakers use the system, which requires a language-in-use approach that looks at naturalistic data and also tests out different uses of classes (cf. the methods used by Franjieh this volume and Singer 2016). In sections 5.7–5.8 I use Mawng as an example to argue that in order to understand the flexibility of a system we need to move beyond a noun-centred approach to nominal classification. This will get us closer to an integrated approach to nominal classification that moves beyond the gender/classifier dichotomy. However, it will mean that we need to focus our efforts on those nominal classification systems analysed only quite recently, for which adequate materials are available.

3 See section 5.7.2 for more evidence for this claim.
5.2 Locating the Mawng language

Mawng is an Indigenous Australian language spoken in the Northern Territory in northern Australia (Figure 5.1). It is a member of the Iwaidjan language family. Mawng is one of eighteen Australian Indigenous languages that are still considered ‘strong’ in that it is still being passed on to substantial numbers of children. There are estimated to be 300–500 Mawng speakers and most children of speakers are still learning the language in the home. Most Mawng speakers grow up speaking more than one Indigenous language, and all adults also speak English. The largest group of Mawng speakers live at Warruwi Community, South Goulburn Island, where most of the fieldwork this chapter is based on was carried out. Mawng speakers also live in other remote Indigenous communities such as Oenpelli and Maningrida and nearby towns such as the regional capital Darwin.4

5.3 Gender agreement in Mawng: A brief overview

Mawng’s gender system is a typical gender system from a morphological perspective. We see plenty of gender agreement in Mawng, to the extent that it could almost be called ‘alliterative’ (Corbett 2006b). There are five genders and a good way to illustrate their differences is to look at how certain bound nominals5 change their meaning depending on their gender agreement prefix. The modifier -lijap ‘small’ is shown in (1). It can be used as a referring expression on its own or can be used as a modifier in combination with other nominals.

(1) Example of a bound noun in Mawng: -lijap ‘small’6

| Masculine (M) | i-lijap | ‘small Masculine entity, e.g. a male human’ |
| Feminine (F) | ninya-lijap7 | ‘small Feminine entity, e.g. a female human’ |
| Land gender (LL) | wu-lijap | ‘small Land gender entity, e.g. a small creek’ |
| Vegetation gender (VE) | ma-lijap | ‘small Vegetation gender entity, e.g. a small stick’ |
| Edible gender (ED) | awu-lijap | ‘small Edible gender entity, e.g. a small yam’ |

4 See Singer and Harris (2016) and chapter 1 of Singer (2016, 2006) for more background on the language and its speakers.
5 The term ‘nominal’ has a particular use in accounts of Australian Indigenous languages. I use the term to include noun-like elements which behave much like nouns but might not be considered ‘nouns’ in other languages such as free pronouns, demonstrative pronouns, modifiers, quantifiers, and certain adverbials. The term nominal covers most common elements of noun phrases but excludes articles and nominalized clauses.
6 Glossing conventions: The pronominal prefixes of transitive verbs are glossed by giving the features of the subject followed by the object, separated by a forward slash (/). The ordering of the subject and object in the gloss does not indicate the ordering of the underlying subject and object affixes. Articles are glossed simply with their gender or number. Free pronouns are glossed just with their person, number, and gender features.
7 This form is irregular, the expected form is niny + lijap → niny-lijap.
Figure 5.1 Map showing area traditionally associated with the Mawng language.
Bound nominals in Mawng constitute a relatively small set; most Mawng nominals do not take gender markers at all. The bound nominals include some but not all of these kinds of nominals: modifiers like -lijap, body parts like -mawurr 'arm', and the numeral -rakap 'one'. These bound nouns do not have an inherent gender; those which are used as modifiers agree in gender with their heads (which may not be overtly expressed). Bound nouns which are body parts agree with their possessors or 'wholes'. The remaining Mawng nouns can be said to have inherent gender. That is, they are typically associated with a particular kind of gender agreement. Example (2) illustrates the extensive gender agreement in Mawng.

(2) Iñy-ura-n-ka jita juk-apa
3F-GO1-PP-HITH F DEM.PROX.F-EMPH1
ngarrunnga-lakajpu-kpu-n la k-anga-ma-ø ngaralk.
3NOM/1PLEX-ASK-KRDP-NP CONJ PRS-3NOM/3LL-GET-NP language:LL
'She (a linguist) came here and was asking us if she could record some language.'
AD Text 11 001

In example (2) we can see that there is gender agreement in the first verb iny-ura-n-ka 'she came here' with the Feminine subject jita jukapa 'this (f)'. This can be attributed to the noun warramumpik 'woman (f)' or to an association between the semantic field 'female human' with Feminine gender. There is also agreement within the referring expression jita jukapa 'this one (female)' as both the article jita and the demonstrative jukapa are forms that are specific for Feminine gender. At the end of the second line of example (2) we see our first noun ngaralk 'language, tongue (LL)'. The verb -ma 'lit: get, here: record' has agreement that indexes ngaralk 'language' as its Land gender object. In example (2) the pronominal prefixes of both transitive verbs show non-Masculine (NONM) agreement as they agree with the female human subject. The non-Masculine category, glossed 'NONM' covers the remaining four genders, apart from Masculine.

Example (2) shows that Mawng has extensive marking of gender agreement. However, gender agreement in Mawng is limited to third person reference. Agreement with first or second persons only indicates person and number, never gender. There are also some forms which have limited gender agreement, agreeing only with two categories: Masculine vs non-Masculine, as we saw in (2).

Table 5.1 lists forms in Mawng that show gender agreement and their limitations.

---

8 This code indicates which text this example is taken from and the utterance number in that text. The examples in this chapter also appear in Singer (2016). Appendix 6 in Singer (2016) lists these codes and gives information on the file and on how to find it in the PARADISEC archive (Pacific and Regional Archive for Digital Sources in Endangered Cultures, <http://www.paradisec.org.au/>).
The classical perspective: The demographics of Mawng genders

One of the aims of this chapter is to argue that the same nominal classification system can look quite different from different perspectives. I will contrast two perspectives on Mawng gender that I call the classical perspective and the language-in-use perspective. The classical perspective relies on the typical resources we have for less well-documented languages. The main resource we have for most languages consists of a dictionary which lists nouns, their genders, and their definitions. The other resource we have for many languages is a grammar which contains many context-free sentences. This classical perspective bases its analysis on a grammar and a dictionary and it is the only view available to us of most nominal classification systems without doing fieldwork ourselves or closely collaborating with a field linguist. However, fuller accounts of diverse nominal classification systems are becoming available, which investigate aspects of the systems that are only evident through analysing language-in-use (for example, see Seifart 2005; Senft 1996; Singer 2016; Wilkins 2000).

The other perspective expands the classical view with an analysis of language-in-use. Senft (1996) points out that the kinds of analysis you can do with a list of nouns and how they are classified are limited. For example, take the task of placing Mawng’s gender system in the possibility space mapped out by the canonical gender approach.

5.4 The classical perspective 1: The demographics of Mawng genders

One of the aims of this chapter is to argue that the same nominal classification system can look quite different from different perspectives. I will contrast two perspectives on Mawng gender that I call the classical perspective and the language-in-use perspective. The classical perspective relies on the typical resources we have for less well-documented languages. The main resource we have for most languages consists of a dictionary which lists nouns, their genders, and their definitions. The other resource we have for many languages is a grammar which contains many context-free sentences. This classical perspective bases its analysis on a grammar and a dictionary and it is the only view available to us of most nominal classification systems without doing fieldwork ourselves or closely collaborating with a field linguist. However, fuller accounts of diverse nominal classification systems are becoming available, which investigate aspects of the systems that are only evident through analysing language-in-use (for example, see Seifart 2005; Senft 1996; Singer 2016; Wilkins 2000).

The other perspective expands the classical view with an analysis of language-in-use. Senft (1996) points out that the kinds of analysis you can do with a list of nouns and how they are classified are limited. For example, take the task of placing Mawng’s gender system in the possibility space mapped out by the canonical gender approach.

---

Note: Where there is allomorphy, the abbreviation for the gender appears, instead of the form.
We would have quite different results if we only took the classical view, compared to a language-in-use view as we will see in later sections of this chapter. Another way to look at the consequences of the two different perspectives is to compare recent work on Mawng (Singer 2016) with an earlier grammar of Mawng (Capell and Hinch 1970). The grammar by Capell and Hinch is recognized as a good-quality grammar for its era. However it was the result of a collaboration between a self-taught linguist (Hinch) who knew the language and an academic linguist (Capell) who had little experience working on the language. Somehow in this process the more unexpected aspects of Mawng gender agreement were overlooked. This is at least partly because a classical view of the language was maintained.

A recent experience brought home the difference between the two perspectives. A colleague who had not previously worked on the Mawng language, was annotating Mawng narratives. To do this he had to analyse the grammatical function of each noun phrase in its clause. Much of this work involved trying to map participants referred to by general kinds of expressions such as nukapa ‘this (Masculine)’ to verbal gender agreement for core arguments. During one meeting, this colleague mentioned that Land gender was clearly the largest gender, because there were a lot of referring expressions with Land gender agreement in the texts. At the time I had the impression that Land gender and Masculine gender were the largest genders in Mawng and that they were of roughly equal size and importance. In fact, if we count the nouns in the Mawng dictionary (Singer et al. 2015), as shown in Figure 5.2, Masculine accounts for more than half of the nouns. From this perspective Land gender is closer in size to the other three genders, than to Masculine.

Figure 5.2 illustrates that the allocation of nouns into genders in Mawng is very unbalanced. Masculine is far and away the largest gender, and while Land and Vegetation genders have a substantial membership, the membership of Feminine and Edible gender is quite small. Based on Figure 5.2, we might expect Masculine gender to be a large miscellaneous class, which has a weak semantic basis, and we would expect Masculine agreement to predominate in texts. In fact this is not an accurate prediction because noun–gender associations are not the only motivations for gender agreement in Mawng discourse. Systems of nominal classification are quite often seen as systems in which nouns are categorized into classes, according to principles of semantic assignment. However, gender in Mawng is an invaluable resource that is used in a range of ways to construct meaning in discourse. The categorization of nouns into classes that we see in Figure 5.2 is a kind of byproduct of how gender is used. Gender agreement in discourse can reflect associations between genders and semantic fields or different construals of referents in context, as well as typical noun–gender associations.

The colleague was applying a GRAID analysis (Haig and Schnell 2014, 2016) for studying discourse in Mawng. Some of the counts based on this analysis are presented in section 5.7.2.
The reason that noun–gender associations are not that important in Mawng discourse is that nouns are relatively rare in naturalistic discourse. Work by Garde (2013) on the neighbouring language Bininj Kunwok argues that there is an areal tendency for referent minimization. This means that speakers tend to minimize the use of referring expressions and even when an overt referring expression is used, it is more likely to be a semantically general form, such as a demonstrative, rather than a noun (see also Garde 2008). If we look at example (2) from Mawng shown earlier, there are no nouns until the third clause. The female linguist, who has already been identified as a discourse topic in the text, is referred to by gender agreement in three words: there is Feminine subject agreement in the verb -ura 'go1', and also Feminine agreement in the article jita 'F' and the demonstrative juka-pa 'this (DEM.PROX.F-EMPH1)'. The only noun in example (2) is the Land gender noun ngaralk 'language' which is the object of the last clause. Further examples of referent minimization from Mawng are discussed in section 5.7.

Cross-linguistically, we find that Mawng is not unusual, as we will discuss more in section 5.7.2. In a great many languages around the world, nouns are relatively
infrequent in naturalistic discourse (Bickel 2003; Stoll and Bickel 2008). Where nouns do occur, they are more likely to refer to inanimate referents that are not subjects, like the noun ngaralk ‘language’ in (2) (du Bois 1987; Haig and Schnell 2016). In Mawng, human referents are more likely to be expressed using verb agreement alone or through semantically general forms such as free pronouns and demonstratives. Meaning is constructed in Mawng primarily through the interpretation of verbal gender agreement and how it is likely to combine with available verb senses (Singer 2012).

5.5 The classical perspective II: Semantics of noun–gender associations

Examining the Mawng dictionary, we find that all five genders have a strong semantic basis. This is illustrated in Figure 5.3 which was generated from the electronic dictionary file, in which each sense of a noun is tagged with at least one semantic field.

The semantic fields in Figure 5.3 are adapted from Voegelin and Voegelin (1957). Inspecting Figure 5.3 we can see that some genders have a strong association with a unique semantic field. For example, the category ‘topography’ is large in Land gender but not elsewhere. However most of the more significant semantic fields are split across at least two genders. For example, animals are found in both Masculine and Feminine gender. Material culture is well represented by both Masculine and Vegetation gender. In addition, plants are a large semantic field in both Vegetation and Edible gender. The large size of Masculine gender means it has more semantic fields than the others. However, the fact that most Masculine nouns fall into clear-cut semantic fields reveals that Masculine is not a ‘leftover’ or miscellaneous class.

The two main semantic principles underlying gender categorization in Mawng are animacy and human gender. As Dahl (2000) points out, the importance of animacy, particularly humanness, in gender systems is often left unstated. In Mawng, the role of animacy is particularly clear as animates are only found in Masculine and Feminine genders. There are no terms such as the German noun Mädchen ‘girl’ which can refer to humans and agree with one of the inanimate genders: Land, Vegetation, or Edible gender. Mawng is slightly unusual in having five genders, as there are many more non-Pama Nyungan languages with a four-gender system. It possible that the Edible gender was carved out of an original larger Vegetation gender. Edible gender contains the edible parts of plants only. Examples include the fruit of fruit trees or the tubers of yam plants. Nearby languages with a four-gender system categorize plants and plant foods into a single gender.

11 The list of semantic domains used in the Mawng dictionary is very similar to the list available at <http://www.anu.edu.au/linguistics/nash/aust/domains.html#beyond>. That list is a version of Voegelin and Voegelin (1957) adapted for use with Australian languages by Kenneth Hale, Jane Simpson, and David Nash.
Figure 5.3 Counts of nouns in the Mawng dictionary by semantic field (1872 semantic field tags from 1254 noun headwords)

Note: Semantic fields with less than sixteen nouns are not represented in this diagram.

Figure 5.4 gives a more detailed breakdown of the main semantic fields associated with each gender.

Figure 5.4 shows how some semantic fields are split across more than one gender. For example, most birds are Masculine gender but a significant number are Feminine gender. This makes sense from a mythological perspective as in numerous Dreamtime myths particular bird species are specifically male or female. They are beings who are like humans and may marry and have children. Those human body part nouns which are not bound nouns are spread over four genders, excluding Feminine gender.

The Mawng dictionary has 3,422 words, 1,985 of which are nouns and 1,254 of these are common nouns with inherent gender, for which a gender has been recorded.
The number of semantic field annotations in Figure 5.3, however, is 1872. The first reason for this is that many nouns have more than one sense listed, and the different senses are often tagged with different semantic fields. In addition, a single noun sense may be tagged with more than one semantic field. For example, the noun *alangunta* has one sense ‘cheesefruit’. This fruit has been annotated with two codes: E.3 ‘plant foods’ and K.5 ‘medicine’ as it is used both as a food and as a cure for the common cold.

Figure 5.3 shows that the nouns in the Mawng dictionary are not exactly balanced across semantic fields. It is worth keeping this in mind while viewing Figure 5.3. Large numbers of plant and animal terms have been recorded which may reflect a focus on ethnobiological research by some linguists working on the dictionary. The dictionary could also be seen by some contributors as a tool in language maintenance and documentation of pre-contact traditions. So we see a lot of words relating to traditional material culture such as fish net types and spear types, most of which are no longer used. The extensive entries for body parts, disease, and other medical
terminology can be attributed to linguist Heather Hewett (née Hinch). She initiated the dictionary and worked on it for two decades but also worked as a missionary nurse at Warruwi Community.

5.6 Canonical gender and the classical perspective on Mawng gender

Section 5.4 outlined the main forms of gender agreement in Mawng, and section 5.5 looked at the semantic tendencies behind noun–gender associations. In this section the information presented in sections 5.4 and 5.5 is used to evaluate Mawng gender using the canonical gender approach. Canonical gender as detailed by Corbett and Fedden (2016) describes an idealized gender system through one principle and three criteria. Canonical gender is a kind of artificial benchmark used to compare actual gender systems with one another. Canonical gender is not intended to represent a possible gender system. In section 5.9 it will be argued that a completely canonical gender system will not be found in natural language. In this chapter we will discuss how Mawng gender sits along the axes of comparison proposed by Corbett and Fedden (2016), from a classical perspective. Then in section 5.8, we will examine Mawng gender from a language-in-use perspective, and see if the view from there differs, in terms of where Mawng is located with respect to canonical gender.

The Canonical Gender Principle states that 'In a canonical gender system, each noun has a single gender value' (Corbett and Fedden 2016: 503). From the classical view of Mawng gender, the system does allocate a single gender value to each noun.12 So this principle is marked ‘Yes’ under ‘classical perspective’ in Table 5.2.

Criterion 1 requires that ‘Canonical gender values match agreement classes’ (Corbett and Fedden 2016: 505). This criterion is closely tied to the Canonical Gender Principle and earlier work on canonical agreement (Corbett 2006b). It is clear in Mawng that gender values match agreement classes fairly well. There may be some syncretism, as we saw in Table 5.1, but it is not of a kind that brings into question the postulated five agreement classes. Criterion 1 as elaborated in Corbett and Fedden (2016: 505–6) states that canonical gender has completely predictable gender agreement. This appears to be the case for Mawng, based on the small amount of data we have examined so far in this chapter, so there is a ‘Yes’ next to Criterion 1 in Table 5.2. However, once we look more closely we will find that variation in agreement patterns creates a lack of predictability, as discussed in section 5.8.

Criterion 2 is phrased, ‘In a canonical gender system the gender of a noun is constant across all domains in which a given language shows agreement’ (Corbett and Fedden 2016: 517). This criterion refers to different morphosyntactic domains

---

12 Although even from the classical perspective some exceptions are noted: for example, nouns such as warranyngiw ‘child’ that can occur with either Masculine or Feminine agreement, depending on the gender of the child.
and compares, for example, gender agreement in the verb to gender agreement within the noun phrase. From the classical perspective, agreement seems to be constant across all domains so it is rated here simply with a ‘Yes’. However, this is because without a language-in-use perspective, we may not come across the rarer examples in which there is a mismatch in gender agreement across domains, like those discussed in section 5.8.

Criterion 3 is also an important one, which will be discussed further in section 5.9: ‘In a canonical gender assignment system, the gender of a noun can be read unambiguously off its lexical entry’ (Corbett and Fedden 2016: 520). This means, surprisingly, that for canonical gender the system of gender assignment is 100% predictable based on the meaning of a noun. So in a canonical gender system the system has 100% semantic transparency. If gender assignment of some nouns is predictable from morphological or phonological properties, or not predictable at all, the system is less canonical.

As we saw in section 5.5, there is a high degree of semantic transparency in terms of how nouns are categorized into genders in Mawng. Yet the gender of nouns is by no means completely predictable. While genders are motivated for at least 80% of nouns, they are probably not predictable for more than 60% of nouns. A number of semantic fields straddle more than one gender, as we saw earlier. Birds may be either Masculine or Feminine and although this classification is motivated by their roles in traditional myths it is not predictable from the semantics of the bird names themselves. Corbett and Fedden (2016) define semantic transparency according to information that is accessible in the lexical entry for a noun. However, we cannot say that Mawng mythology is information that is accessible in the lexical entry.

There are many nouns in Mawng whose gender associations are neither predictable nor motivated. Take, for example, the distribution of free human body part

| Table 5.2 The classical perspective on Mawng gender and canonical gender |
|-----------------------------|-----------------------------|
| **Canonical gender**       | **Classical perspective on Mawng gender** |
| Principle                  | Yes                         |
| Criterion 1                 | Yes                         |
| Criterion 2                 | Yes                         |
| Criterion 3                 | High level of semantically transparency |
nouns across four genders. Why should *ngaralk* ‘ear’ be Vegetation gender, *wanji* ‘head’ be Land gender, and *maruturt* ‘heart’ be Edible gender? Mawng is quite canonical with respect to Criterion 3 but not completely. In section 5.9 the tension between what is dictated for flexibility on the one hand and semantic transparency on the other for canonical gender is discussed. We might expect that a certain degree of semantic transparency favours flexibility in use of nominal classification systems, which creates a tension between the different axes of canonicity for gender.

### 5.7 The distribution of nouns in Mawng narratives and its implications for the role of gender agreement in discourse

The classical account of Mawng gender presented so far has considered noun–gender associations. We looked at how gender agreement is expressed on various parts of speech (section 5.3), how Masculine gender is associated with more nouns than all the other genders put together (section 5.4), and how nouns can be grouped into semantic fields according to their reference (section 5.5). Then in section 5.6 we looked at how Mawng’s gender system would be categorized in the canonical gender approach, when looked at from this classical perspective. In this section, we move beyond the classical view, with its focus on noun–gender associations, to a more dynamic view of Mawng gender. We will see how Mawng gender agreement is used in naturalistic discourse, often in the absence of any nouns and sometimes even in ways that seem to conflict with noun–gender associations. From this language-in-use perspective, we see that noun–gender associations cannot explain all gender agreement observed in discourse.

As mentioned in section 5.7, the noun, while clearly present as a part of speech in Mawng, is not the omnipresent meaning-bearing element it is in languages such as English. In fact, linguists exert themselves to explain cases of zero anaphora in English because omission of a referring expression is the exception (e.g. Fillmore 1986). Whereas in Australian languages the onus is on linguists to explain why referring expressions appear where and when they do, particularly those for core arguments (Schultze-Berndt and Simard 2012; Simard 2010). In Mawng, as in many Australian languages, we find that every referring expression receives a pitch accent (Fletcher and Butcher 2014; Fletcher et al. 2015). Compare this pattern with English in which referring expressions only receive a pitch accent if they are in focus, for example, or the last element of an intonational phrase (Fletcher et al. 2015). The fact that every referring expression receives at least one pitch accent suggests that they are much more prominent in Mawng discourse than they are in English. The scarcity of Mawng nouns allows gender agreement to play a much more important role in the construction of meaning in discourse. In 5.7.1, I look at the frequency of gender-agreeing forms across seven Mawng narratives and in sections 5.7.2–5.7.3 I look at the frequency of nouns in the same corpus.
5.7.1 The prevalence of gender-agreeing forms in narrative texts

In Figure 5.2 (section 5.4), a pie graph was presented which showed the number of noun–gender associations of various kinds in the Mawng dictionary. Figure 5.2 showed that over half of all nouns are associated with Masculine gender. However, it was also noted that a linguist annotating Mawng texts had the impression that Land gender was the largest gender. The reason for this discrepancy seems to be that the frequency of noun–gender associations in the dictionary differs substantially from the frequency of gender-agreeing forms in texts. In Figure 5.5 the frequency of gender-agreeing forms in seven narratives totaling 784 clauses is shown.

As we saw in section 5.3, gender agreement in Mawng is found in the verb and within the noun phrase as well as in nominalized clauses, demonstratives, and pronouns. Comparing Figure 5.5 with Figure 5.2, we can see that Masculine and Feminine gender are of roughly equal proportion in both. That means that the proportion of types in the dictionary database is equivalent to the proportion of tokens in the corpus. However, when we look at the other three genders we find that the proportions are quite different. There are a much larger proportion of Land-gender-agreeing forms in the corpus than there are Land gender nouns in the dictionary. The proportions of Vegetation gender and Edible gender also do not match across Figures 5.2 and 5.5. The proportions are much smaller in Figure 5.5 than in Figure 5.2. These discrepancies show that the distribution of genders in the dictionary does not predict the frequency of gender-agreeing forms in discourse. One reason for this could be that noun–gender associations are not the only factor influencing gender agreement in discourse. This is discussed further in section 5.8. The following two sections, sections 5.7.2 and 5.7.3 look at the frequency of nouns in the seven narratives. The scarcer nouns are in Mawng discourse, the less likely it is that they alone determine gender agreement.
5.7.2 The scarcity of referring expressions in Mawng

Recent work in grammatical typology has shown that there are many languages around the world in which referring expressions are scarce in discourse. Bickel’s work on referential density showed that this is not only the case in languages with rich agreement such as Mawng (Bickel 2003; Stoll and Bickel 2008). Many East Asian languages which lack rich agreement also have a low proportion of referring expressions in discourse. However, languages with rich agreement like Mawng remain a group of languages in which we can reasonably expect a lower density of referring expressions in discourse (Seifart 2011; Seifart et al. 2010). In this section we analyse the frequency of referring expressions in the seven narratives discussed in section 5.7.1, using the GRAID annotation system (Haig and Schnell 2014). The GRAID annotation system has been used to compare the use of referring expressions in narratives across typologically and genetically diverse languages (Haig and Schnell 2016).

The seven texts in the corpus are all narratives; most are monologues but some include a little interaction with a second speaker. The total number of clauses in the corpus is 784, which includes subordinate clauses. The GRAID discourse annotation method is used because it pays attention to animacy, argument status, and whether a given referring expression contains any nouns. The GRAID method distinguishes pronouns from lexical noun phrases. Lexical noun phrases are those referring expressions which contain nouns whereas the contrasting category ‘pronouns’ covers referring expressions that consist only of a free pronoun or a demonstrative (Haig and Schnell 2014). The percentage of core arguments expressed as lexical noun phrases is shown in Table 5.3.

How the figures for Table 5.3 were calculated is best explained by an example. For the first cell in the table, the number of lexical noun phrases in transitive subject function (A), we divide the total number of lexical noun phrases found in transitive subject function (22) by the total possible number of referring expressions in transitive subject function (307) and multiply the result by 100 to convert it to a percentage.14 Table 5.3 shows that the number of lexical noun phrases is relatively low across all functions as we would expect, but higher for intransitive subject (S) and object (O) function. The proportion of pronouns does not make up for the shortfall, leaving transitive subjects as those most likely to be expressed by zero anaphora. In Table 5.4, we look only at the human referents in the data, to see if there are any differences.

What Table 5.4 shows is simply that there are even fewer overt referring expressions for human referents and far more zero anaphors. In Mawng, human referents are much more likely to be expressed using verbal agreement alone or through semantically general forms such as free pronouns and demonstratives than through lexical noun phrases. This is a very common pattern which is observed, it seems, universally in

---

14 So, for example, the first cell in the table = Lex.NP.A/(LexNP.A+Pronoun.A+zero.A).
language (du Bois 1987; Haig and Schnell 2016; Meakins 2015). Haig and Schnell (2016) compare the frequency of lexical noun phrases across nineteen corpora of very diverse languages. Comparing the frequency of lexical noun phrases in Mawng narratives with their figures shows that referring expressions in Mawng are relatively scarce but Mawng is by no means exceptional. Similar frequencies of lexical noun phrases were found in Haig and Schnell’s corpus for Yagua and Spanish. The counts for this chapter were made using the same criteria as the counts in Haig and Schnell.15

### Table 5.4 Referring expressions used for core arguments in the seven-text corpus (human referents only)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>S</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical noun phrases</td>
<td>2%</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>Pronouns</td>
<td>7%</td>
<td>5%</td>
<td>12%</td>
</tr>
<tr>
<td>Zero noun phrases</td>
<td>91%</td>
<td>93%</td>
<td>84%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

5.7.3 **Noun scarcity in practice**

Section 5.7.2 showed that Mawng noun phrases have a scarce distribution in discourse and that this is not unusual cross-linguistically. It is not obligatory to use a referring expression to express every core argument. Even when a referring expression is used, for a human referent it is more likely that a pronoun will be used than an

---

15 In consultation with Stefan Schnell.
expression containing a noun, as Table 5.4 shows. This pattern is also well attested cross-linguistically (Haig and Schnell 2016). To refer to certain semantic fields, nouns are rarely used in Mawng. For example, although nouns exist that can be used to refer to speech, stories, and knowledge such as ngaralk (LL) ‘tongue, speech, story, knowledge’, wularri (LL) ‘discipline, (traditional) knowledge’, mayali (VE) ‘wisdom, traditional knowledge’, these nouns are only rarely used. Instead these kinds of arguments are simply implied or are expressed by a gender-agreeing form such as a demonstrative or relative clause. For example, in (3) a story is referred to using the modifier nominal -lurtpuj ‘short’, a demonstrative pronoun takapa ‘distal, Land gender’, and a relative clause, indicated by the square brackets.

(3) Nga-mi-ny nuwurr-apa wu-lurtpuj-apa taka-pa
1sg-say-pp 2pl.OBL-EMPH1 LL-SHORT-EMPH1 DEM.D.LL-EMPH1
[tangung-purr−n]RelC
LL 1sg/3LL-know-NP
‘I told you a short story that I remember.’

The article ta at the start of the relative clause in (3) indicates that the following material is functioning like a noun phrase. The fact that it is Land gender suggests that it refers to an entity of Land gender, namely the object of the relative clause verb—wurr-ru ‘think, know, remember’. Main verbs are never preceded by the article so the presence of an article immediately before a verb flags a relative clause construction or nominalization of another type. The absence of the noun ngaralk ‘story, tongue, speech, language’ from the sentence in (3) is typical for utterances that refer to speech, stories, or knowledge. Example (3) comes at the end of a storytelling session, so the topic ‘story’ is accessible in some ways in the broader context. However, referents can be deemed recoverable in a very wide range of contexts, far beyond what is usually considered ‘given information’.

It is not only abstract referents like ‘story’ in example (3) that tend to be expressed in ways that do not involve nouns. The utterance in (4) introduces two new objects into the narrative, without using nouns to refer to them. The text is an account of the arrival of the first missionary at Warruwi, Goulburn Island.

(4) Something arriwana-ma-ø
Eng 1PL.IN/3M,FUT-GET-NP
kunuka ja apani-lati-ki-n.
what M 3M/3ED,FUT-have.inside-KRDP-NP
‘We’ll get something (a boat) that can transport (food).’ pajawali 2017

---

16 See Singer (2010) for more examples of references to speech, story, and knowledge.
In example (4) the quoted speech introduces two new referents, a boat and a shipment of food, through verbal agreement, without any referring expressions at all corresponding to these referents. What turns out to be a boat is referred to as the third person Masculine object of the verb -ma 'get' in the first clause in (4). This boat is referred to by the noun kapala 'boat (M)' four lines after (4). The other underspecified new referent in (4) is waliq 'food (ED)' which is the object of the second verb -lati 'have inside' in (4). The food is referred to using the noun waliq 'food (ED)' six lines later. It is likely that hearers would be able to reconstruct the referent 'food' more easily as Edible gender has much fewer semantic fields than Masculine gender. It is also likely that many audiences would have already have heard the story of the first missionary before, and therefore have no trouble working out what the two underspecified referents of (4) are. However, what we see here is still indicative of a stylistic tendency to use nouns sparingly in Mawng discourse.

Cross-linguistic work suggests that there are areal patterns in stylistic tendencies to avoid noun phrases or use them more liberally (Haig and Schnell 2016; Stoll and Bickel 2008). Garde (2008, 2013) looks extensively at the use of referring expressions in conversation in Bininj Kunwok, a language of a different language family to Mawng that has traditionally been spoken in the same area as Mawng. He identifies a tendency for 'reference minimisation' and finds that even when a hearer indicates that they are having problems resolving a reference, the speaker may still not provide enough information for the hearer to resolve that reference. This also seems to be the case in Mawng. If a storyteller uses minimal references like those to a boat and food in (4), hearers will not necessarily interrupt and complain the storyteller should be more specific. These kind of oblique references seem to be more tolerated in Mawng even if they do leave things unclear for the hearer.

To what extent is it valid to analyse the gender agreement in examples (3) and (4) as being controlled by some absent noun? Certainly, to defend a noun-centred approach to nominal classification we could argue that deep in speaker’s minds is a lexical database of noun–gender associations, just like the dictionary of Mawng analysed in section 5.4. This underlying knowledge could be argued to motivate speakers’ use of agreement, even in the absence of overt nouns in the discourse. However, this may not be the best way of describing the use of gender agreement in languages such as Mawng. In the following sections I argue that a more integrated approach to nominal classification systems that overcomes the traditional gender/classifier dichotomy is only possible if we are not only concerned with nouns, but rather encompass the full range of reference management resources that are available (Merlan, Roberts, and Rumsey 1997). In the following section I propose that there is a direct association between genders and semantic fields in Mawng, not mediated by noun–gender associations. The evidence for this is that semantic field–gender associations can override noun–gender associations.
5.8 Departures from the Canonical Gender Principle in Mawng language in use

In this section we will look at contexts in which verbal gender agreement can be different from that predicted by the noun–gender associations discussed in sections 5.4–5.6. There are also other kinds of contexts in which gender agreement in Mawng discourse can deviate from that predicted by noun–gender associations. These are discussed in Singer (2016). In this section we will look at how Mawng departs from the two elements of canonical gender that specifically mention nouns: the Canonical Gender Principle and the Canonical Gender Criterion 2. The Canonical Gender Criterion 2 is effectively a narrower version of the Principle. The Principle states that each noun should only be associated with one gender and Criterion 2 specifies that the gender agreement for a noun should not vary across syntactic domains. We do not always find one-to-one mapping of nouns to genders in Mawng. Mawng gender agreement departs from both the Principle and Criterion 2 at times. Where there is a mismatch between verbal gender agreement and noun–gender associations, Mawng gender agreement also departs from Criterion 2 as gender agreement need not be consistent across all syntactic domains. Verbal gender agreement in Mawng can differ from NP-internal agreement for the same referent, as illustrated in (5).


‘She understands a little bit of knowledge.’

In example (5) there is Vegetation gender agreement within the noun phrase which is what we would expect as the noun–gender association of the noun mayali ‘(traditional) knowledge, wisdom’ is with Vegetation gender. The modifier nominal -lijap ‘small’ in the noun phrase agrees in gender with the noun mayali by taking the Vegetation gender form of its obligatory prefix. However, the verb -atpi takes Land gender object agreement, as it usually does when it means ‘understand’. The other sense of this verb is the more general meaning ‘have, hold’ which occurs with a range of kinds of object gender agreement. However, when the verb means ‘understand, grasp (knowledge)’ it usually has Land gender object agreement. This is not completely unexpected as the semantic field of knowledge is also associated with Land gender (see section 5.5). What we see in example (5) is that verbal gender agreement is not always predictable from noun–gender associations. Rather, verbal gender agreement may index the semantic field of a referent, even when an overt noun is also used to refer to that referent, which has a different gender. This is the case in

---

17 Singer (2016) also contains many more examples of the kind presented in sections 5.7 and 5.8 of this chapter, for readers who are interested.
example (5). Instead of indexing the gender associated with the noun *mayali* (Vegetation gender), the gender agreement in the verb indexes the semantic field ‘knowledge’ associated with Land gender (see Figure 5.4).

The variation from expected gender agreement that we see in (5) is not dissimilar to the phenomenon described as ‘hybrid nouns’ in the literature on European language gender. Corbett and Fedden (2016) use the term ‘lexical hybrids’ and cite the German word *Mädchen* ‘girl’, the most well-known example. It has neuter gender within the noun phrase but anaphoric pronouns referring to girls may take feminine agreement. Like the noun *mayali* ‘knowledge (Vegetation gender)’, the noun *Mädchen* ‘girl (neuter)’ is an example of a mismatch between noun–gender association and semantic field–gender association. In the case of *mayali* it fits into the semantic field of ‘knowledge’ which is associated with Land gender. However, the noun–gender association of *mayali* is with Vegetation gender. Similarly, in German the noun *Mädchen* ‘girl (neuter)’ fits into the semantic field ‘female humans’ which is associated with feminine gender in German, but the noun is associated with neuter gender. This reflects morphological rules for gender assignment in German, which are as important as semantic considerations. Nouns with the diminutive suffix -*chen* usually have neuter gender agreement.

Zubin and Köpcke (2009) and Köpcke, Panther, and Zubin (2010) discuss the various contextual factors that influence whether anaphoric pronouns referring to the noun *Mädchen* ‘girl’ take neuter forms or feminine forms. They find patterns specific to particular genres of writing. For example, the use of neuter gender agreement with *Mädchen* ‘girl’ throughout is a stylistic feature of fairy tales, even new ones. There are also tendencies that relate to how the ‘girl’ is being presented; as more child-like and therefore non-gendered (neuter) or as more adult and gendered (feminine). They give an example of an article about an adult female model where anaphoric agreement is feminine throughout. Whereas a newspaper article about a female child who is injured, uses neuter agreement throughout.

Previous accounts of hybrid nouns such as *Mädchen* ‘girl’ suggested that the alternative feminine gender agreement patterns reflected the ‘natural’ gender of the female referent. Zubin and Köpcke (2009) propose the ‘pragmatic projection principle’ whereby gender is projected from individual noun–gender associations to broader associations between genders and semantic fields. It is interesting that even in a language like German where there appears to be little semantic basis to gender classification, semantic fields are relevant. Zubin and Köpcke (2009) suggest that these projected semantic fields have expanded over time to become more important motivators for gender agreement in German (p. 249).

Like the German hybrid noun *Mädchen* ‘girl’, what the clash between noun–gender associations and semantic field–gender associations does in Mawng is create variation and reduce the predictability of gender agreement. This increased variation is what increases the flexibility of a nominal classification system. Note that in
addition to apparent mismatches in agreement as we saw in (5) the noun *mayali* ‘knowledge’ can also occur with the expected Vegetation verbal agreement as in (6).

(6) Marrik anm-atpi-ø mira mata mayali.  
NEG 2SG/3VE-understand-IRR EMPH2 VE knowledge (VE)  
‘You can’t understand that knowledge.’

Comparing examples (5) and (6) we can see that there are two possible options for verbal gender agreement when the verb *-atpi* ‘understand’ combines with the object *mayali* ‘knowledge’. The verb *-atpi* ‘understand’ can also have more predictable agreement when it has a Land gender object noun such as *wulatpiyi* in (7).

(7) K-anny-atpi-ø [wu-murra wulatpiyi]NP.  
PRS-3M/3LL-understand-NP LL-much meaning (LL)  
‘He has a lot of knowledge.’

In fact, in most naturalistic examples, the object of *-atpi* is not referred to by a noun phrase at all. Example (8) has three tokens of *-atpi* but the object is not expressed by a noun phrase for any of these tokens.

(8) Ta nuwurri yara ta wurrwurr-awk marrik kurrung-atpi-ng.  
ART 2PL some ART new-COLL NEG 2PL/3LL-understand-PC  
‘Now some of you, the new people, you can’t understand.’

Ngarri-pa arrung-atpi-ø-wi parang-apa  
1PL.EX-EMPH1 1PL.IN/3LL-understand-NP-only finish-EMPH1  
la arrungpan-atpi-ø pirij.19  
CONJ 1PL.IN/3LL.FUT-HAVE-NP drop  
‘Only we can understand but we’re going to pass it (knowledge) on (to you).’

This section has shown how competition between noun–gender associations and gender–semantic field associations can lead to variation in gender agreement patterns. We can now return to the comparison of Mawng gender with canonical gender, begun in section 5.6.20 In that section we took a classical perspective on Mawng gender. In section 5.7 and in this section we have extended that perspective by looking at the use of gender and referring expressions in discourse, to add a

18 This sentence was explained as an example of what older people might say to younger people when they are arguing.
19 The coverb *pirij* forms a complex verb construction together with the verb *-atpi* ‘have, hold’ in this example. Coverb constructions are presented by displaying the meaning of the coverb construction as a whole under the coverb. The inflecting verb has a gloss in capitals which attempts to give an underlying meaning for that verb across all coverb constructions it has been recorded in.
20 See section 5.6 for a definition of the Canonical Gender Principle and the three criteria.
Table 5.5 Perspectives on Mawng gender compared

<table>
<thead>
<tr>
<th>Canonical gender</th>
<th>Classical perspective on Mawng gender</th>
<th>Language-in-use perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principle</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Criteria 1</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Criteria 2</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Criteria 3</td>
<td>High level of semantic transparency</td>
<td>High level of semantic transparency</td>
</tr>
</tbody>
</table>

Beyond the classifier/gender dichotomy

5.9 Flexibility in the use of nominal classification systems as a key feature of the canonical gender approach

Typologies of nominal classification that make a primary division between classifiers and genders often also mention flexibility. However, they identify flexibility of classification as a characteristic of classifier systems. The associations between classifiers and nouns do not always motivate what is seen in discourse; rather, speakers can be creative in how they categorize referents. In the Papuan language Imonda, for example, a coconut can be referred to by different classifiers while it is in different states of being processed: from being a ‘picked’ coconut, to having its husk removed, its shell broken, and then once its shell is removed (Seiler 1985:132, quoted in Merlan, Roberts, and Rumsey 1997). Examples of these shifts or what Corbett and Fedden (2016) call ‘recategorisation’ are well known for many ‘classifier languages’. For instance, Reid (1997) gives an example from the Australian language Ngan’gityemerri in which a boomerang is referred to by the THING classifier rather than the usual WEAPON classifier because it is used as a doorstop. Wilkins (2000) discusses how nouns can be combined with novel classifiers for humorous effect in Mpartnwe Arrernte classifier constructions, giving the example Artwe Toyota (MAN Toyota) ‘Man of the Toyota dreaming’, in which the vehicle make Toyota is used in place of the name of a sacred site.
This flexibility is also found in otherwise apparently gender-like systems like Mawng. We saw some examples of the flexibility of the Mawng system in section 5.8 and more extensive evidence is provided in Singer (2012, 2016). Many systems which have been categorized as classifier systems, on the other hand, give speakers very little flexibility. Franjieh (this volume) shows how the system of possessive classifiers in North Ambrym has, unexpectedly, very little flexibility at all. After testing out numerous combinations of classifiers and nouns, Franjieh concludes that the combinations are quite inflexible. He shows that possessive classifiers are found in most of the languages of Vanuatu, but the degree of flexibility with which they can be used varies greatly. In Australia we also find that systems which appear quite similar syntactically can vary greatly in their flexibility. Gaby (2006) describes the classifier system in Kuuk Thaayorre, an Australian language of the Pama-Nyungan family like Mparntwe Arrernte, mentioned in the preceding paragraph. The Kuuk Thaayorre classifier constructions look superficially like the very flexible constructions in Mparntwe Arrernte. However, Gaby’s observations on the use of Kuuk Thaayorre classifier constructions lead her to suggest that all combinations of classifiers and nouns are ‘stored as a separate entry in the mental lexicon rather than produced on the fly’ (Gaby 2006: 84).

The kind of flexibility discussed in this section is often referred to as ‘cross-classification’ in typologies of nominal classification. It is said to be a key feature of classifier systems that is lacking in a prototypical gender system (Dixon 1982; Grinevald 2000). Corbett and Fedden (2016) also discuss this flexibility, locating it within the broader phenomenon of ‘recategorisation’. From a canonical gender approach, flexibility is always non-canonical as it entails that gender agreement cannot be predicted purely from noun–gender associations. After a brief discussion of the flexibility of gender agreement in the language Savosavo, Corbett and Fedden (2016) make the observation below:

In a language like Savosavo we see that the less the extent to which gender agreement is fully determined by unique gender values, the greater its role in constructing meaning.

(Corbett and Fedden 2016: 526)

Corbett and Fedden (2016) point to work on Mawng and some other languages as supporting their observation. With respect to Mawng, I have argued that both the flexibility in the use of the genders and their semantic transparency allow the system to be used to construct meaning in discourse, in ways which would not otherwise be possible. Prior to Corbett and Fedden (2016), earlier typological approaches distinguished genders from classifiers on the assumption that a number of properties of nominal classification systems correlate. For example, it was assumed that if a system had a greater number of classes, they were more likely to have a stronger semantic basis, therefore more likely to be used in a flexible manner, therefore more likely to contribute to the construction of meaning in discourse. These long chains of
correlations made for rather messy typologies. And these messy typologies were rooted in the assumption that genders and classifiers form discrete categories, an assumption which does not seem to be borne out by recent data.

The canonical gender approach replaces the gender/classifier dichotomy with a concern with flexibility. While agreement is rather uncontroversially held to be a defining feature of gender systems, what is new in the canonical gender approach is the centrality of flexibility. Canonical gender is a rigid system of nominal classification in which the class used is always determined by a noun and there is never any question as to what class is used in any context, in any genre, etc. Compared to earlier approaches, the canonical gender approach is more elegant. However, it raises a number of questions about the messy correlations assumed by earlier typologies. These questions concern three different properties of nominal classification systems, which are by no means straightforward to evaluate and compare:

1. Flexibility in the use of a nominal classification system.
2. Functions of a nominal classification system, including the ways it can be used to construct meaning in discourse.
3. The semantic transparency of a nominal classification system.

The quote from Corbett and Fedden (2016) given earlier predicts a correlation between points 1 and 2 in (9). They mention point 3, semantic transparency, separately, in another section of their article, without making a link to properties 1 and 2. Flexibility of the kind found in Mawng, however, seems logically to depend on a degree of semantic transparency. It is difficult to imagine how Mawng gender could be used creatively, if there were not a high degree of semantic transparency in all five genders. With the three potentially interdependent properties of nominal classification in (9) we seem to have come full circle, back to a messy chain of correlations. If a system is flexible, it is possible for it to have a broader range of functions, but it can only be more flexible if it has greater semantic transparency. However, the canonical gender approach makes progress by putting flexibility first and foremost. This means that we do not have to assume a correlation between semantic transparency and flexibility, which would be disproven by cases such as North Ambrym and Kuuk Thaayorre that, as mentioned earlier, have a high degree of semantic transparency but a low degree of flexibility. Without being deterministic we can hypothesize that greater semantic transparency tends to favour greater flexibility in the use of a nominal classification system, rather than assuming a correlation.

Seifart (forthcoming) defines semantically transparent assignment as ‘the matching of the descriptive content of the noun class morphology with the semantics of the classified noun’. However, more than one description can match, as illustrated by the case of the Miraña term for turtle he discusses, which was used earlier with the ‘round’ classifier and is now used with the ‘animate’ classifier. Seifart (2005) explores how we might compare the semantic transparency of different nominal classification
systems. While it is uncontroversial to claim that the German gender system is less semantically transparent than Mawng’s there is a large middle ground where it can be difficult to make a call. There is a need for a more rigorous approach to measuring semantic transparency. It may be possible to operationalize the variable, by evaluating to what extent nouns—in particular genders—form natural classes, as explored in section 5.5 for Mawng. The difference between predictable semantic assignment and motivated semantic assignment mentioned will be relevant to how semantic transparency is measured.

5.10 Extending the insights from the canonical gender approach to forge a more integrated typology of nominal classification

Systems of nominal classification are exceedingly diverse. The various classes in a nominal classification system can be marked in many different ways, grammatical agreement being just one quite grammaticalized option. In addition, criteria for categorization into classes can encompass a wide range of semantic contrasts, even if most are based around an animacy contrast, masculine/feminine contrast, or a shape contrast (Croft 1994; Dahl 2000). If we look at nominal classification quite broadly, it is challenging to create a typology that encompasses all the diversity. Gender systems are often more easily defined so that we find much more debate about the boundaries of the typological category ‘nominal classification’ than about the boundary of the category ‘gender systems’ (Hellwig 2007; Passer 2014). Mawng has a system of five genders which in many ways conforms to the general understanding of what gender should be. However, a definition of gender that defines gender as a property of nouns will never be that useful for explaining how gender is used in Mawng. In discourse we find that gender agreement can be motivated by a range of associations that often override typical noun–gender associations. For example, there are associations between genders and semantic domains. Once we map out the entire galaxy of nominal classification strategies as fully as possible, it may be possible to identify common kinds of gender systems that are relatively time-stable systems and develop from common grammaticalization pathways.

The aim of canonical gender theory is to give linguists the tools to compare nominal classification systems with one another in a rigorous way. In order to achieve this, it aims to do away with the distinction between genders and classifiers. The gender/classifier dichotomy that muddies the water for typologies of nominal classification is tied to the Eurocentric history of work on nominal classification. As Kilarski (2013) shows, the Eurocentric bias of work on nominal classification is not simply a data issue: that we still have far more data for gender systems in European languages than nominal classification systems of any type elsewhere. The problem is also one of theory; the giants whose shoulders we stand on worked mainly on gender in European languages. The other well-developed area of research in nominal
classification is East Asian classifiers. These systems are very different, so give the false impression that gender and classifier systems are easily distinguished. In fact, there are also a vast number of systems which do not fall neatly into either the category ‘gender’ or ‘classifier’.

Like Corbett and Fedden (2016), this chapter takes the view that the gender/classifier dichotomy stands in the way of better typologies of nominal classification. However, this chapter also argues that in order to move beyond the gender/classifier dichotomy, we need to broaden our focus from nouns to other constructs that nominal classification classes may associate with. I have demonstrated that noun–gender associations give us only part of the picture of how nominal classification in Mawng works, by contrasting what I called the ‘classical view’ of Mawng gender in sections 5.4–5.6 with the language-in-use view in sections 5.7–5.8. I propose ‘semantic fields’ as a way forward, as there is evidence that classes are associated with semantic fields as well as nouns in Mawng and other languages. For example, as discussed in section 5.8, Zubin and Köpcke (2009) argue for the relevance of semantic fields in German. The importance of noun–gender associations relative to other gender associations is yet another axis of variation in nominal classification systems. Clearly, in German, semantic fields do not yet play as significant a role as in Mawng.

Corbett and Fedden’s (2016) Canonical Gender Criteria give us three axes of comparison. However, we need more in order to map out the world’s nominal classification systems in ways that reveal significant correlations and patterns of grammaticalization. To really engage with the diversity of the world’s nominal classification systems, these systems must be contextualized in terms of the roles they play in constructing meaning in discourse. Many systems, particularly those that are less well-known, are likely to play a crucial role not only in reference but also predication, as is the case in Mawng (Singer 2012). By examining the use of Mawng’s nominal classification system in this chapter, we have seen how a system works in which nouns do not play such a central role.

Broadening our view of nominal classification systems, so that we see them as more than just noun-classifying devices, is essential to gaining greater insight into why there are so many different nominal classification systems around the world. The prominence of nouns in current typologies is fundamentally misleading. Seifart (forthcoming) demonstrates that the degree of transparency in nominal classification systems is something that changes over time. He shows that the classification of nouns into classes is a contingent, dynamic effect of how nominal classification is used, rather than something that is basic to the system. This chapter presents an argument for the relevance of language-in-use data to the typology of nominal classification. The classical view of nominal classification systems emphasizes the influence of noun–gender associations on gender agreement in discourse. A language-in-use view emphasizes, as does Seifart (forthcoming), how the use of a nominal classification system structures categorization of nouns and referents into classes. The argument
made here is that a language-in-use account of a nominal classification system provides us with a better understanding of how the system fits into typological space than a classical account.

Acknowledgements

Warruwi Community has given their support to this research and I particularly thank all the Mawng speakers who have been involved. Peter Hurst worked on the GRAID annotation and Stefan Schnell provided advice on the GRAID annotation and analysis as well as useful comments on this study. Thanks to Greville Corbett and two anonymous reviewers for their comments on this chapter. This research was supported by an Australian Research Council DECRA postdoctoral fellowship (DE140100232), Faculty Research Grant (Arts Faculty, University of Melbourne), and the Research Unit for Indigenous Language (School of Languages and Linguistics, University of Melbourne). The chapter was revised while I was a visitor to the Crossroads project and research associate of the Department of Linguistics, SOAS (University of London), and I thank them for their support. The visit was supported by several travel grants: a Dyason fellowship (University of Melbourne), a Manchester University–Melbourne University cooperation grant (School of Languages and Linguistics, University of Melbourne), and a Fritz Thyssen Foundation travel fellowship (Germany).