Part II

Perspectives from Syntactic Description and Theory

5

Features in categorization, or a new look at an old problem

Keith Plaster and Maria Polinsky

5.1 Introduction[1]

A large body of linguistic literature is concerned with syntagmatic properties of linguistic features: how are features distributed, what is the domain of their operation, what happens when several features compete? All of these issues notwithstanding, it is also important to keep an eye on the paradigmatic aspects of features, and one of the primary questions here has to do with the extent to which linguistic features can and should be taken for granted. This question concerns any set of possible linguistic features, and there may very well be no uniform answer. This chapter examines the content of gender[2] features possible in human language.

We would like to start with an analogy with phonetics. Humans are capable of producing (and distinguishing) a significant number of sounds, but phonological systems around the world are typically constrained by general principles of intelligibility, distinctiveness, or optimal design. This sets in motion a system of checks and balances that allows for significant recurrence of phonetic features across the world's languages.

If we now turn to the category of grammatical gender, gender features can, in principle, be of any kind – nouns in a language can be categorized on the basis of various properties of their denotations, from function to size to colour. But, just as one does not find vowel systems consisting solely of /i/, /y/, and /u/, we do not find gender categorizations based on colour, curly shape, or density. Why? Is it simply because linguists have not been looking closely enough or in the right places, or is it because such systems are impossible?

[1] We are grateful to Claire Bowern, Greville Corbett, Mark Harvey, Jay Jasanoff, Anna Kibort, Andrew Nevins, Jeremy Rau, the audience at the Features Workshop at University College London, and two anonymous reviewers for their helpful comments. All errors are our responsibility.

[2] Noun class and gender are different terms denoting the same concept (Corbett 1991: 1); 'class' and 'gender' will be used interchangeably in this chapter.

This chapter is an attempt to revisit this general issue and to argue that, at least with regard to gender classifications, the inventory of features employed by different languages is quite small. The limitations imposed on such an inventory are ultimately due to learnability constraints and perceptual salience. To support this point of view, we will examine a particularly difficult case of noun classification.

Much work on noun class/gender systems has focused on the role of semantics in the assignment of nouns to classes. Sometimes the semantics responsible for noun class assignment is rather straightforward and based on general, superordinate categories such as 'human', 'animate', or 'female', which are known to be foundational in developmental psychology. Sometimes, however, the kinds of semantic categories used to account for noun class composition are of a very culture-specific, complex nature (for an overview of the relevant literature, see Craig 1986; Corbett 1991: 307–324).

Before appealing to complex semantic and cultural information as the basis for noun classification, it is helpful to gauge the extent to which this information is available to speakers, and in particular to those trying to get a toehold in the system. Young language learners do not yet have access to much of the abstract and culture-specific information possessed by adults, and may have substantial difficulties learning a noun classification system based on such knowledge. The idea that gender systems are learned early is strongly supported by the acquisition of gender in Bantu, where the system that needs to be learned is quite complex, with languages continuing the canonical Bantu noun class system generally containing six noun classes paired for singular and plural and another set of approximately six unpaired classes (Katamba 2003). For a number of Bantu languages, it has been shown that gender is learned by the beginning of the third year of life and that both production and comprehension of gender are practically error-free (Kunene 1979; Suzman 1991, 1996; Kapust 1998; Demuth 1988, 1992, 2000, 2003; Deen 2005, 2006; Idiata 2005).

Outside of Bantu, the complex gender systems of German, French, or Russian are also learned early, and this learning is manifested both in comprehension and production (Mills 1986; Karmiloff-Smith 1979; Gvozdev 1961; Kempe et al. 2003; Polinsky 2008).[3] The knowledge of gender is revealed in multiple ways – from the use of correct declension classes correlating with

[3] In each of these languages, small subclasses of nouns resist correct gender assignment until an older age, usually 4 or 5. For example, in Russian, certain neuter nouns and feminines ending in a palatalized consonant continue to pose a problem for learners until age 5 (Gvozdev 1961: 442; Smoczyńska 1985: 644; Polinsky 2007). In German, some plural endings (in *-e*), whose gender assignment is not always straightforward, remain problematic for children up to age 4;0 (Mills 1985, 1986).

genders (cf. Gerken et al. 2005, where English-speaking 17-month-olds were able to distinguish grammatical from ungrammatical Russian words after approximately two minutes of exposure to a Russian gender paradigm) to the correct use of gender agreement (Deen 2006; Kempe and Brooks 2001; Karmiloff-Smith 1979, among others). Crucially, the learning of all these systems relies on simple conceptual cues (animacy, mobility) and salient formal features. For example, the acquisition of Slavic gender is facilitated by the frequent use of diminutives (Kempe and Brooks 2001; Ševa et al. 2007); the diminutives provide consistent *formal* cues, which are further reinforced by regular agreement patterns. Overall, the generalization is that young learners pay attention to the combination of salient formal features of a word *in conjunction* with agreement.

At the same time as young learners are mastering the gender systems of various languages, their abstract thinking lags behind their grammatical competence. Numerous studies show that children in the second year of age are just beginning to develop incipient symbolic reference (Gallistel 1990; DeLoache 1995). Moreover, children in their second and third year of life map words and objects referentially, not associatively (Gallistel 1990; Preissler and Carey 2004; Clark 1993; Gelman and Bloom 2000; Gelman 2000, 2003; Mandler 2004). This in turn suggests that outside of the most frequent items, whose grammatical information is likely to be learned by rote as word islands (cf. Tomasello 1992), the young learner acquiring gender is sensitive to what he or she knows how to identify best – word segmentation, stressed segments, syllabic division, or, in other words, salient formal properties of language (e.g. Hudson Kam and Newport 2005; Weiss and Newport 2006).

To the extent that language change is driven by children acquiring language, we should expect the kind of information to which children have access to play a significant role in the structure and development of a language. If so, gender systems which are based on complex semantic cues and strong meaningful analogies should present an insurmountable obstacle to learning. Either such systems should be learned much later in life and/or prove unstable, or it may be that their inherent structure is amenable to a different analysis, one that does not rely on complex semantic features.

In this chapter, we will take up a particularly well-known case of exotic noun classification, that in Dyirbal, a moribund Pama-Nyungan language of Australia. We argue that the development and operation of Dyirbal gender can be accounted for on the basis of a small number of semantic features andthe morphophonemic similarities of items and their type/token frequency without recourse to complicated semantics. If our proposal is on the right track, this arguably sophisticated gender system does not constitute a

counterexample to the general conception that grammatical gender is acquired on the basis of straightforward linguistic cues.

The remainder of the chapter is structured as follows. In Section 5.2, we introduce the composition of the Dyirbal nominal lexicon. Section 5.3 reviews the existing analyses of Dyirbal noun classification. Section 5.4 is a brief survey of our proposal concerning the diachronic development of the Dyirbal genders, and Section 5.5 presents a proposal concerning the synchronic composition of Dyirbal genders such that it could be accessible to a young learner. Section 5.6 takes our reanalysed system to the next step, where this system is further simplified and undermined by loss of agreement cues. Section 5.7 presents our conclusions.

5.2 Data

Dyirbal has four noun classes (genders), which are manifested in agreement with the demonstrative ('noun marker', Dixon 1972). Demonstratives also mark case and differentiate between proximal, medial, and distal deixis. For medial deixis, the gender markers are shown in Table 5.1.

Moving on to the composition of noun classes, Dixon (1972) presents the following list of concepts associated with each class and suggests that such a heterogeneous list might indicate that the noun classification lacks any principled basis. However, he also notes that speakers have little variation in class assignment and immediately assign loanwords to a class. The breakdown of concepts by classes is shown in Table 5.2.

This table shows a complicated system, which poses a serious challenge to any learnability model. Even if we set young learners aside, the storage and access of such a classification system nevertheless seem challenging for an adult speaker. Therefore, accounting for such a system is particularly important for our understanding of noun categorization in natural language. A plausible account should try to explain the principles underlying the system and also make predictions for the classification of new words entering the

TABLE 5.1 Dyirbal noun class markers, medial deixis

	Absolutive	Ergative	Dative	Genitive
Class I	*bayi*	*baŋgul*	*bagul*	*baŋul*
Class II	*balan*	*baŋgun*	*bagun*	*baŋun*
Class III	*balam*	*baŋgum*	*bagum*	—
Class IV	*bala*	*baŋgu*	*bagu*	*baŋu*

TABLE 5.2 Distribution of concepts across noun classes in Dyirbal (Dixon 1972: 307)

I (*bayi*)	II (*balan*)	III (*balam*)	IV (*bala*)
men	women		parts of the body
kangaroos	bandicoots		meat
possums	dog		
bats	platypus, echidna		
most snakes	some snakes		
most fishes	some fishes		
some birds	most birds		
most insects	firefly, scorpion, crickets hairy mary grub anything connected with fire or water		bees and honey[a]
moon	sun and stars		wind
storms, rainbow			yamsticks
boomerangs	shields		
some spears	some spears		some spears
	some trees	all trees with edible fruit	most trees
			grass, mud, stones
			most noises, language

[a] While Dixon (1972) identified *girnyjal* 'honey' as belonging to class IV, Dixon (1984) identifies the noun as belonging to class III.

lexicon. Dixon and several authors after him have proposed a purely semantic account of the Dyirbal noun class system, to which we turn in Section 5.3.

5.3 The semantic account of gender assignment in Dyirbal

The existing accounts of Dyirbal noun classes, proposed by Dixon (1972), Lakoff (1987), Mylne (1995), and Harvey (1997: 19–24), all emphasize the conceptual underpinnings of the classification.

To account for the distribution of concepts across the four noun classes in Dyirbal (Table 5.2), Dixon suggests that class membership in Dyirbal is best

explained through the interaction of some basic (core) concepts associated with the various classes and a set of overriding rules. The core concepts are (Dixon 1972: 308):

(1) Core semantic concepts in Dyirbal noun classification

Class I:	animacy, (human) masculinity
Class II:	(human) femininity; water; fire; fighting[4]
Class III:	edible plants (nonflesh food)[5]
Class IV:	everything else (residue)

Although in theory the association between core concepts and noun classes could be quite straightforward, in actuality nouns are assigned to classes subject to a set of conceptually based rules. These rules override the general principles of class assignment listed in (1); in other words, the principles in (1) represent the default conditions, whereas the overriding rules allow for a more fine-grained noun-to-class network, based on more specific conditions.

At this point, we will present the overriding rules followed by examples of their application.[6]

(2) *Rule 1 (Myth-or-belief)*: If a noun has characteristic X (on the basis of which its class membership would be expected to be decided) but is, through belief or myth, associated with characteristic Y, then generally it will belong to the class corresponding to Y, not to X.

For example, most birds are in class II because they are thought to be spirits of dead women. However, willy wagtails, which as birds should be in class II, are instead in class I because they are believed to be mythical men. Several other birds are also assigned to class I based on such mythical connections (Dixon 1972: 308).

[4] In a different account, Mylne (1995) suggests that the main opposition is in terms of potency and benign/malign power. This still leaves the account driven by complicated semantics, a feature we are arguing against here.

[5] Such a narrowly defined class is rather surprising given the three other classes which are very inclusive. But note that in Australian traditional culture, 'children must learn from the earliest age to be able to classify plants (and indeed anything else they are liable to put into their mouths as babies)' (Mylne 1995: 390). See also Walsh (1993).

[6] Dixon (1972: 308–310) proposes the first two rules; rule three is proposed by Lakoff (1987: 93) on the basis of Dixon's analysis.

(3) *Rule 2 (Domain-of-experience)*: If there is a basic domain of experience associated with A, then it is natural for entities in that domain to be in the same category as A (Lakoff 1987: 93).

To illustrate, plants tend to be in the elsewhere class in Dyirbal (class IV), but plants that have edible fruit are in class III with the fruit they yield.

(4) *Rule 3 (Important-property)*: If a subset of referents has some particular important property that the rest of the set does not have, then the members of that subset may be assigned to a different class from the rest of the set to 'mark' that particular important property. In Dyirbal, the important property is often [+harmful].

For example, most types of fish are in class I; however, harmful ones (stonefish, toadfish) are in class II. Similarly, harmful (stinging) trees are in class II, with the harmful fish, but all other trees without edible fruit are in class IV.

Following Dixon's analysis, Lakoff (1987) proposes a description of the Dyirbal gender system based on the notion of a radial category. Within a radial category, a particular element that has most of the defining characteristics of that category serves as the prototype. Other elements are assimilated to the prototype on the basis of their perceived resemblance to the prototype, but they do not have to actually share the criterial features of that prototype. The more peripheral members are linked to the prototype through other members, and these links can be motivated by certain principles. Taken together, the members of a category thus form a radial structure, with the most representative, or prototypical, members located at the centre, and with less representative outliers clustered around this hub.

In the case of Dyirbal class II (containing 'women, fire, and dangerous things'), the core is human females (Lakoff 1987: 100–101). The links to the core are achieved in the following manner: the sun is a female deity and is married to the moon (which is in class I, as a male). The sting of the hairy mary grub hurts like sunburn; thus, it is linked to the sun. All items related to fire (fire, matches, and pipe) are linked to the sun as well. The firefly is linked to fire. Stars and light are associated with and linked to fire, and through it to the sun.

Birds are in the female class because they are spirits of dead women. (And some birds are in class I because they are spirits of dead men.) Crickets in myth are 'old ladies' and they are in the female class. Since women are related via myth to the sun, and the sun via domain-of-experience is related to fire, women and fire end up in one class. Since fire is dangerous, other dangerous things should be in the same class. This brings in water and fighting.

So far, we have applied the myth-or-belief rule and the domain-of-experience rule. Now comes the important-property rule (4): because something is 'harmful', it is placed in a separate class from all other members of its subset. Thus, most fish are in class I, but harmful fish are in class II. In other words, noun class distinctions can be used to underscore differences in some critical characteristic (see Zubin and Köpcke 1986 for a similar principle in the analysis of German gender assignment).

Mylne (1995) modifies Lakoff's account to make it more culture-specific and sensitive to the semiotics of the traditional aboriginal society. However, despite these modifications the analysis remains largely semantically driven and expands the notion of the prototype. Thus, theoretically, it is not very different from Lakoff's account. And in another drive towards a more culture-specific account, Harvey (1997: 24) proposes an ontology in which natural concepts such as [human], [force of nature], [edible] are associated with particular environments (air, tree, water, ground), with further classifications built on such associations. While the details of his proposal may differ from the others discussed here, the spirit is very similar: identify a salient concept and build a radial category by various semiotic links to this concept.

If the rules described above applied satisfactorily and if the radial category account were comprehensive and had predictive power, then one could conclude the investigation of the Dyirbal noun classification. However, the rules do not apply in any systematic way and, as they are, seem to act more as after-the-fact generalizations than operational principles (see also footnote 12). The radial-category account fails to motivate the links in an unambiguous and predictive manner. Most importantly, if links between members of a radial category require specific cultural knowledge and often have to be explained at length, it raises the crucial question of learnability: how does a young learner acquire all these links and relations? In our view, the opacity of links inside the presumed radial category suggests that an alternative account would be desirable.

Crucially, in their discussion of the Dyirbal genders, the previous authors rarely mention the actual lexical items occurring in Dyirbal – instead, they deal only with semantic concepts and the referents of the lexical items. In other words, they have not tried to find any connection between the form of the Dyirbal words appearing in a certain class before turning to the words' semantics. Given that learners are known to be very sensitive to formal cues from infancy (Saffran et al. 1996a, b; Newport and Aslin 2000, 2004; Hudson Kam and Newport 2005), and that young learners tend to regularize inconsistent input based on segmental information, it is important to give

formal cues the credit they deserve before turning to complicated semantic cues. In addition, if formal cues were shown to fail then the categorizing power of conceptual structure could be validated even more forcefully than in the earlier accounts presented here.

In what follows we will attempt to explain Dyirbal noun classification differently, and our results will be compatible with the very general features that are found in gender systems all over the world. In a nutshell, we will portray Dyirbal as a much less exotic language as far as notional motivation for gender is concerned. In particular, we will argue for a combination of formal and semantic cues in Dyirbal gender assignment, bringing its gender classification much closer to the familiar systems around the globe. Before discussing these issues in Section 5.5, we will briefly address the diachrony of gender classification in Dyirbal.

5.4 Diachronic origins of Dyirbal noun classes

The category of gender in Pama-Nyungan is rare, although there are numerous instances of languages with classifiers. In Plaster and Polinsky (2007), we review evidence from closely and more remotely related Pama-Nyungan languages and show that they have comparable classifier systems which find correspondences in the composition of Dyirbal noun classes.

In addition to the evidence for classifier systems in languages closely related to Dyirbal, Dyirbal itself provides evidence of its former classifier system. Dixon and Koch (1996: 44) note that the Dyirbal gender system is a relatively recent development, as confirmed by the relatively infrequent use of noun class markers in Dyirbal song poetry (see Plaster and Polinsky 2007 for the details of this evidence).

The original classifier system of Dyirbal likely contained many more classifiers that can no longer be identified as synchronic semantic classes. Gender systems can develop from classifier systems through the collapse of a larger number of classifiers into a smaller number of genders, with generic nouns such as 'woman', 'man', and 'animal' serving as class cores (Corbett 1991: 311–312, 317). During the development of the Dyirbal classifier system into the present gender system, different classifiers merged to create each noun class. Although classifiers themselves are semantically based, the composition of the genders resulting from such a merger are not necessarily driven by the semantics of the classifiers; although one may expect classifiers for clearly semantically related concepts (such as 'edible' and 'potable') to merge, semantically unrelated classifiers may merge based on shared formal

features or for no currently clear reason (the semantic motivation may have become opaque).

In Plaster and Polinsky (2007), we hypothesize that correspondences between the classifier classes and the resulting noun classes in Dyirbal are as follows:

(5) Dyirbal classifier classes > Dyirbal noun classes

Class I	[+male], [+edible animate]
Class II	[+female], [+bird], [+fire], [+fresh water], [+stinging]
Class III	[+edible non-animate]
Class IV	everything else

The merger of several of the classes we have identified can be motivated straightforwardly based on the semantic characteristics involved. As we will discuss below, [+male] and [+female] appear to be the semantic cores of classes I and II respectively, and the distinction between animate and inanimate nouns in Dyirbal is strong. All animate nouns fall within classes I and II; while classes I and II contain some inanimate nouns, no animate noun appears in class III or class IV. Accordingly, it appears that animacy was likely a driving factor in the development of the noun class system.

Focusing on *animate* nouns, a merger of the [+male] class with the [+edible animate] class led to the exclusion of those animate nouns that are not edible. Since only two classes, I and II, include animate nouns, this entailed that animates which were not [+edible animate] ended up in class II: any other animals that were not identified as 'edible', including dogs, flies, grasshoppers, spiders, and worms, were placed into class II. Similarly, animates capable of inflicting a harmful sting were placed in class II due to their animacy, and not because they were identified with human females. As non-edible animals, they could not be placed in the other animate class, class I; this left only class II available, and the entire set of [+stinging] items was drawn into class II.

However birds are a problem for this proposal: they are included within the scope of the [+edible animate] classifier, for example, in the closely related language Yidiny, but they are in class II in Dyirbal rather than the expected class I. This may indicate that the merger of birds with human females occurred at an early stage during the collapse of the classifier system, possibly due to formal attraction or analogy, or the early identification of birds with the spirits of dead human females.

One possible attractor for such a merger of birds with human females into a single class is the feminine suffix *-gan*, which is found in a number of

languages on the east coast of Australia (see Dixon 1972: 12–13 for references), including Banjalang (Crowley 1978: 37). For example, Dyirbal *jarrugan* 'scrub hen' appears to be derived from the Dyirbal generic noun or classifier corresponding to the Yidiny classifier *jarru*, possibly consisting of *jarru-* 'bird' and the feminine suffix *-gan*. Identification of the ending *gan* with the feminine suffix *-gan* could easily have led to the form being attracted into the class for human females. Other bird names in our sample also contain formal features identified with class II.[7]

In assigning gender to a new noun, a speaker may classify it on the basis of formal similarity with familiar nouns (rather than semantics), especially if the initial or final segments of the noun resemble a pattern that the speaker recognizes as associated with a particular class. In addition to the salient suffix *-gan*, a crucial role in formal analogy belongs to stressed segments. The role of stressed segments in the division of nouns into classes is well known from familiar Indo-European languages; for instance, Latin or French gender assignment can be successfully explained by appealing to stressed endings (Tucker et al. 1977; Corbett 1991: Chapters 2 and 3; Polinsky and van Everbroeck 2003; Lyster 2006). In Dyirbal, the stress is invariably on the first syllable, and we find strong evidence indicating that the initial syllable played a role in creating formal analogy that could facilitate class mergers.

For example, *yimalimal* 'welcome swallow' begins with *yi-*, which is strongly correlated with class II and is the initial syllable of *yibi* 'woman'; all animate nouns in our sample with the initial syllable *yi-* are placed in class II.[8] In addition, several bird names begin with *bi-*, which is also strongly correlated with class II membership, likely due to the association with the form *bibi*, which appears in various Pama-Nyungan languages with the meaning 'woman', 'mother' or '(female) breast' (O'Grady 1998). 'Whistle duck', 'large parrot', and 'white ibis' all begin with *bi-*. We will discuss each of these formal features in detail in Section 5.5, but we would like to emphasize that the

[7] Our sample contains 597 Dyirbal nominal forms culled from the following sources: Dixon (1972, 1980, 1982a, b, 1984, 1989, 1990), Dixon and Koch (1996), and Schmidt (1985). In the original lexicon, there are a number of items from the avoidance language (a highly formalized register used for communication across moieties, found in a number of Australian languages), which we did not include in our analysis because avoidance languages are often associated with their special phonology or grammar. In addition, we did not include items that appear only in Dyirbal song poetry. Thus, here we treat only forms from Guwal, the everyday, non-avoidance language. A spreadsheet of the forms used in our sample is available upon request.

[8] As we explain in Section 5.5, although *yirrinyjila* 'dragonfly' begins with *yi-* in the Dyirbal dialect of Dyirbal, the form in the other dialects begins with *wi-*. As a result, in all other dialects the form does not contain the formal feature.

existence and salience of these features was probably the initial motivation for the classification of birds with human females.

Of the *inanimate* classes, the class of edible inanimate items was clearly the most salient, as attested by its status as the sole semantic class found in class III. As with the merger of names for birds and human females, we propose that the merger of fire- and water-related items with the other members of class II was due to existence of formal attractors and not any semantic identification between human females and fire or fresh water. For example, *bugan* 'brush fire' ends in *-gan*, which is identical to the feminine suffix *-gan*. Likewise, *binda* 'waterfall' begins with *bi-*, which we have also identified as a formal feature associated with class II (see Section 5.5).

Formal features allow us to directly account for two of the prime examples for the 'domain-of-experience' rule in (3): *garri* 'hairy mary grub' and *yarra* 'fishing line'. Although we would expect to find *garri* in class I, the default class for animate nouns, it appears in class II. Dixon (1972: 310) explains its classification as due to its semantic association with the sun, noting that 'its sting is said to feel like sunburn'. However, Dixon makes nothing of the exact correspondence between the forms *garri* 'sun' and *garri* 'hairy mary grub', or that this correspondence provides a more direct explanation for the class assignment. Put more directly, we propose that *garri* 'hairy mary grub' is not a class II noun because its sting is similar to the effect of spending too much time in the *garri* 'sun', but because its form is identical to *garri*.

In fact, recognizing that formal attraction led to the placement of both nouns in class II allows us to explain the assignment of *garram* 'garfish' to the same class. Although most fish are assigned to class I, a few fish, including garfish and stonefish, are assigned to class II, which Dixon and Lakoff attribute to the operation of the important-property rule, provided in (4); under their analysis, a subset of fish possess an important property ('harmfulness'), and this property is marked by placing these fish in a different class from other fish. However, while stonefish are, in fact, harmful, garfish are not; no Australian garfish species is known to be dangerous or harmful, although the members of one species are 'described as "pugnacious . . . but are incapable of inflicting anything like a serious wound"' (Mylne 1995: 395). This and certain other assignments of non-human animates to class II led Mylne (1995) to question Dixon and Lakoff's determination that harmfulness underlies the assignment of these non-human animates to class II.

As noted above, we propose that the Dyirbal [+edible animate] classifier class merged with the class of male humans to form class I. Under this scenario, 'garfish' was originally placed in class I as 'edible', while 'stonefish' and 'toadfish' were excluded from class I as not 'edible', just as 'garfish' bears

the 'edible meat' class *minya* in Yidiny, while 'stonefish' and 'toadfish' do not. Accordingly, the assignment of 'garfish' to class II is similarly unexpected.

We propose that the assignment of *garram* 'garfish' to the same class as *garri* 'sun' and *garri* 'hairy mary grub' is due to the phonological shape of the left edge of each word. No other animate noun in our sample begins with *garr-*; all three – the sun being animate by myth – are assigned to class II.[9]

A similar argument can be made for the assignment of *yarra* 'fishing line'. Although, as an inanimate noun, we would expect to find *yarra* in class IV, it is found in class I. This unexpected classification is explained by Dixon as due to the semantic connection between men and fishing. However, the phonological similarity between the actual form for 'man', *yara*, and 'fishing line', *yarra*, more directly explains the class assignment. These forms are identical except for the rhotic in each form. Although /r/ and /rr/ are phonemes in traditional Dyirbal, as shown by the minimal pair in question, the /r/ and /rr/ distinction appears to have been breaking down at the time of Dixon (1972), as shown by consistent dialectal differences in the presence of /r/ or /rr/ word-finally, as in the forms for 'navel', *jujur* in the Dyirbal and Mamu dialects but *jujurr* in the Giramay dialect, and 'urine', *jujar* in the Dyirbal and Mamu dialects but *jujarr* in the Giramay dialect. Thus, we propose that the rhotic distinction was not sufficiently salient to prevent the attraction of *yarra* 'fishing line' into class I on the basis of *yara* 'man'. This proposal receives support from the observation that the distinction between the two phonemes is almost lost in Young People's Dyirbal (YD), the language of the children and grandchildren of Dixon's Dyirbal consultants.[10]

The identification of formal features by speakers may either encourage the reassignment of a noun from its semantically expected class to the class with which the formal feature is associated (as in the case of *garri* 'hairy mary grub', *garram* 'garfish', and *yarra* 'fishing line'), or prevent the reassignment of a noun from a semantically unexpected class – but with which the formal feature is associated – to another, semantically expected class.[11] The result

[9] The Dyirbal word *garambarri* 'young alligator' is the only other animate noun in our sample that begins with the sequence *ga-*, but due to its association with the class I nouns *gujagay* 'alligator' and *maybaja* 'alligator' we would not expect its classification to be affected by *garri* 'sun'.

[10] According to Schmidt (1985), the distinction between these phonemes has been fully lost in YD in words except where the distinction is necessary to prevent homophony, as in the case of *yara* 'man' and *yarra* 'fishing line'; otherwise, YD speakers wavered in their realizations of the /r/ and /rr/ phonemes. The weakening or collapse of rhotic distinctions has also been documented for other dying languages (see Schmidt 1985: 193 for references and discussion).

[11] Which of these two forces is more powerful is an interesting question, but beyond the scope of this chapter.

of either scenario is identical: a noun's class matches the formal feature that it carries.

If the origins that we propose for the noun class system are correct, the language documented by Dixon (1972) represents an intermediate stage in the expected development of the noun class system. The semantic criteria for certain classes have changed over time as speakers in general and child learners in particular were faced with the task of identifying the system's underlying principles, as shown by the shift from [+edible animal] to [+non-human animate] as a component of class I. The difficulty of identifying these semantic principles became even more troublesome once reclassifications began to occur on the basis of formal features. An extremely shallow agreement system, where the demonstratives were the only exponent of gender, added to the difficulty of maintaining gender in Dyirbal (compare this shallow system with the pervasive agreement found in Bantu, Northeast Caucasian, or Indo-European gender languages). If given sufficient time, we would expect the number of exceptional classifications in the Dyirbal system to dwindle, and the vast majority of nouns to be classified simply and straightforwardly on the basis of a small number of core semantic classes and clear formal features.

5.5 Motivating Dyirbal noun classifications synchronically

In accounting for synchronic noun categorization we approach it first and foremost from the standpoint of learnability. When faced with the task of determining which noun class to associate with each noun, a Dyirbal child is not able to draw on sophisticated semantic concepts and connections that either adults (seeking to justify the class associations that they have learned)[12] or linguists (seeking to find an underlying order in the system) may come up with. A child has no inherent (or learned) association of women with dangerous things, contrary to Lakoff's account, or as an 'other' and 'associated with the disruption of harmony of living', contrary to Mylne (1995: 387). Since many of the concepts that Lakoff and Mylne identify as underlying

[12] Lakoff (1987: 100–101) cites the following speech recorded by Schmidt (1985) from a semi-speaker of Dyirbal as evidence for a connection in the minds of Dyirbal speakers between fire and danger, on the one hand, and women, on the other: 'buni [fire] is a lady . . . Woman is a destroyer. 'e destroys anything. A woman is a fire.' However, it is impossible to determine whether the speaker possessed this belief prior to being asked, or if the speaker actually relies on the stated association to classify the items. If asked, even English speakers, who possess a minimal gender system, can posit semantic reasons for ships being referred to as 'she' rather than 'it', while in actuality it is merely a matter of custom.

the Dyirbal noun class system are beyond the scope of young children's understanding, the systems posited by Lakoff and Mylne would be nearly impossible for children to learn. Unless we expect Dyirbal children to have memorized all class associations by rote until such time as they could understand the complex concepts and relations posited by the semantic accounts, the existing explanations of the Dyirbal gender system are difficult to maintain.

While complex mythological associations that require rote learning, abstract knowledge, and vast cultural experience are unlikely for a 2-year-old, developmental psychology shows that children under the age of one are able to differentiate such basic categories as 'human', 'animate', 'male', 'female', and 'mobile' (Gentner and Namy 1999; Namy and Gentner 2002; Kellman and Arterberry 1998; Mandler 2004, among many others). Remarkably, these categories match the basic semantic categories involved in gender assignment across languages (Corbett 1991: 7–30, 82–89), and we can reasonably expect such a core to be present in our system. With nouns that fall outside of those semantic cores, one could expect a child to make errors and to reassign the nouns on the basis of some other salient cues, most likely the phonological form of a word. For example, children should learn quickly that males and most non-human animates belong in class I, females in class II, edible items in class III, and inanimate things in class IV. A child faced with the need to determine which class marker to use with *jirrga* 'eel-spear' and *baŋgay* 'spear' would place them in class IV with the other inanimate nouns that the child knows, rather than class I and II, respectively.

5.5.1 *Semantic core*

When we examine the composition of the noun classes in Dyirbal, we find that classes I through III have a well-defined semantic core. This core is reinforced by the fact that nouns denoting the relevant concepts do not appear elsewhere, so the core is in a sense exclusive to the relevant class. Nouns referring to male humans appear only in class I, while nouns referring to female humans appear only in class II. Consumable, non-beverage items other than meat – a more accurate characterization than 'edible', which we use henceforth for the sake of brevity – appear only in class III. Examples are provided in (6):

(6) a. Class I: [+male]
yara 'man'
gaya 'mother's younger brother'
wirru '(potential) husband'

b. Class II: [+female]
bulgu '(potential) wife'
jugumbil 'woman'
gajin 'girl'
c. Class III: [+edible]
jugur 'wild yam'
gabi 'gabi fig'
wuju 'vegetable/fruit food (generic)'

Loanwords from English are also drawn into the relevant classes on the basis of the same semantic criteria, as shown in (7):

(7) a. Class I
bulijiman ~ *buliman* 'policeman'
waybala 'white man'
b. Class II
mijiji 'white woman (missus)'
c. Class III
binarra 'peanut'
gaygi 'cake'
laymun 'lemon'

Thus, any [+male], [+female], or [+edible] nominal will be assigned to class I, class II, and class III respectively.[13] In addition, the majority of non-human animate beings, such as animals, fish, and insects, appear in class I, and in fact comprise the majority of the class I nouns in our sample, while the vast majority of inanimate nouns are placed in class IV. Accordingly, class I appears to be the default class for non-human animate beings, while class IV is the default class for inanimate nouns.

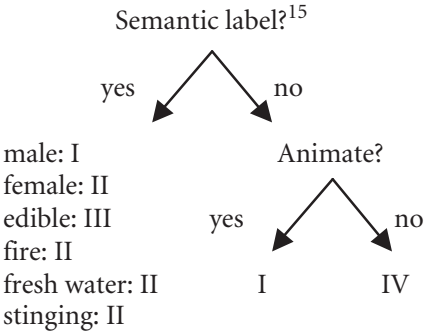
While class III is a limited semantic class, only admitting edible items, and class IV is an elsewhere class, accepting nouns that are not assigned to any of the other three classes,[14] classes I and II contain a variety of nouns that do not fall within the semantic classes identified above. The three smaller semantic classes that we identified in Section 5.4, [+fire], [+fresh water], and [+stinging], account for the classification of certain inanimate nouns to

[13] The exceptionless character of these class assignments leads us to conclude with Dixon (1972) and Lakoff (1987), and against Mylne (1995), that masculine and feminine are core features of class I and II respectively.

[14] Due to its elsewhere nature, class IV is by far the largest: of the 597 nouns that we have been able to establish, about 49% belong to class IV.

class II. A decision tree for determining the assignment of Dyirbal nouns based on the above features is shown in (8):

(8)



While the previous accounts of the Dyirbal gender system sought a synchronic semantic basis for the assignment of every Dyirbal noun to its class, we propose that those nouns whose classification does not follow directly from (8) are classified on the basis of formal features or are remnants of the earlier classifier-based system which have successfully resisted reassignment due to their frequency of use.

5.5.2 *Beyond the core: in search of synchronic motivations*

Recall that classes I and II contain certain nouns whose class assignment is unexpected. In addition to words for humans, classes I and II contain non-human animate nouns as well as various inanimate items, including the words for 'sun', 'moon', and 'rainbow'. Although class I appears to be the default class for non-human animate beings, certain non-human animate beings appear in class II, including most birds and certain fish. If class I is the default class for animates, any animates found in class II require an explanation. Similarly, we must explain any inanimate nouns placed in either of these classes.

As emphasized by Dixon (1972, 1984) and Dixon and Koch (1996), the role of folklore in Dyirbal culture is undeniable, and we agree that folkloric

[15] We separate animacy from the remaining semantic features for two reasons. First, unlike the other semantic features, which could be assigned to nouns not possessing the semantics of the feature (as, for example, in the class assignment of the moon and sun based on their mythological identification as male and female respectively), [+animate] appears not to have been a feature that could be applied to a noun unless its referent truly was animate. In addition, in the event of a conflict between an assigned semantic label and a noun's inherent [+animate] feature, the assigned semantic label determined the noun's class assignment (as in the case of the animals belonging to class II due to their identification as [+female] or [+stinging]).

associations are likely responsible for the class assignment of at least some nouns found in class I or II. The categorization of words denoting celestial objects (stars, planets) is likely due to folklore. As Dixon explains, the moon is the husband of the sun in Dyirbal folklore, and accordingly *gagara* 'moon' is placed in class I and *garri* 'sun' in class II. Similarly, *yamani* 'rainbow' is a man in Dyirbal mythology, and assigned to class I. In addition, according to Dyirbal folklore, birds are the spirits of dead human females, and most bird names are in class II.[16] Similarly, crickets, which were excluded from the class of edible animals that formed class I, are assigned to class II because speakers liken them to 'old ladies'.[17]

Since the telling and retelling of folklore was such a part of Dyirbal society,[18] from an early age children learned the mythological semantic labels associated with these nouns, and this accounts for their gender assignment; it certainly helps in the learning of their gender that these words are few in number and presumably at least some of them are frequent. If certain words were not such an integral part of Dyirbal everyday conversation or storytelling, we would expect them to be susceptible to reassignment to class I, in the case of birds, or to class IV, in the case of inanimate nouns. In fact, these are precisely the reassignments that Schmidt (1985) documented in the language of the children and grandchildren of Dixon's Dyirbal consultants, as we will discuss in Section 5.6.

The situation with the smaller semantic subclasses is similar. To the extent that a child is able to identify that nouns associated with water or fire or that are capable of harmful stinging are placed in class II, the association of these items with class II will remain. Fresh water and fires both played a large role in Dyirbal daily life, and words relating to fresh water and fire do not appear in any other class. As a result, it is easy to believe that children received sufficient input to identify that fresh water- and fire-related items belong in class II.

[16] The names of birds follow mythological associations much less reliably than the names of celestial objects or require more extensive knowledge of the avian world. In addition to birds like willy wagtails, which are placed in class I as mythical men, Dixon mentions that certain other birds (like hawks) go into class I because they eat other birds. Rather than being classified as [+male], we suspect that such bird-eating birds may not bear the [+female] label due to their carnivorous conduct, and as a result are placed in the default class for animates.

[17] Again, the question of chronological order arises: were crickets viewed as 'old ladies' before they fell into class II, or did the connection between crickets and 'old ladies' come about later as an explanation for the class assignment?

[18] We would like to thank Mark Harvey for a very helpful discussion regarding the importance of storytelling and folklore in Australian aboriginal culture.

Similarly, the class of 'stinging' items appears to remain a semantic class relevant to noun class assignment, but one that was breaking down at the time of Dixon (1972). The 'stinging' nouns in our sample include:

(9) Animate nouns

bima	'death adder'
gabul	'forest carpet snake'
gadambal	'mangrove crab'
gumbiyan	'echidna'
jaŋgan	'stonefish'
juruŋun	'toadfish'
malayigarra	'scorpion'
marrigal	'chicken snake'
munilan	'chicken snake'
yunba	'water python'

(10) Inanimate nouns

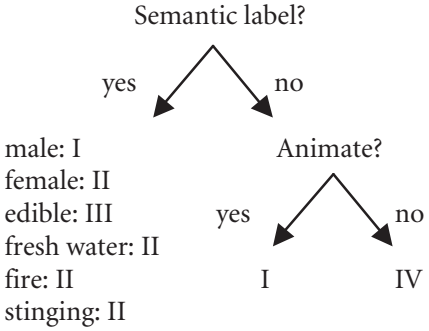
bumbilan	'stinging nettle'
jaŋali	'small stinging tree'
duŋan	'stinging tree'
giyarra	

Although not all of the animates that we have identified as possible members of the [+stinging] class are poisonous, each is capable of causing a harmful stinging by biting, pinching, or otherwise attacking a human. For example, although pythons are not poisonous, the water python is a particularly aggressive snake. Similarly, the mangrove crab can exceed two kilos and possesses powerful claws.

We propose that the relevant semantic feature remains 'capable of inflicting a harmful sting' rather than simply 'dangerous' or 'harmful', as Dixon and Lakoff propose, as shown by assignment of 'harmful' nouns to other classes. Many animals assigned to class I are capable of inflicting harm, including crocodiles, alligators, the brown snake (one of the most dangerous snakes in Australia), and the *bujimburran*, a stinking beetle that squirts fluid (and identified as a 'danger to eyesight') (Dixon 1984: 148). Thus, due to the assignment of all stinging plants and the stinging animate nouns identified in (9) to class II, we posit that [+stinging] is a synchronic classification relevant for class assignment, albeit a weaker classification than the other classifications identified above.

In sum, under the view proposed here, the semantic assignment of gender in Dyirbal was based on the following principles:

(11)



The semantic labels match some of the original classifiers that we hypothesized Dyirbal had; the range of these classifiers is also supported by the evidence from the classifier systems of related languages (see Plaster and Polinsky 2007). The development of larger gender classes resulted in the loss or bleaching of semantic motivation for smaller classes, and gradually resulted in a greater role of formal cues. Had Dyirbal stayed healthy as a language and not undergone attrition and gradual death under the encroachment of English, its four-class system may have developed further, with formal cues as criterial factors. Instead, the system underwent a significant reanalysis in Young People's Dyirbal, which we will discuss in Section 5.6 below. Before turning to this new language, however, we would like to discuss the formal cues that were available in traditional Dyirbal.

5.5.3 *Beyond the core: formal cues in gender assignment*

The formal features responsible for class assignment affect only the animate nouns in our sample. That such a restriction would exist is not surprising given the distribution of nouns across the Dyirbal classes. Animate nouns appear only in classes I or II, and both of these classes consist primarily of animate nouns. Inanimate nouns, on the other hand, appear predominantly in class IV. Upon determining that a noun is animate but not inherently male or female, a child learner knows that the noun will appear in class I or II; as a result, the amount of motivation necessary to pull a class I noun into class II will be less than that required to pull an inanimate noun into class II. While formal features that affect both animate and inanimate nouns may also be present in Dyirbal, we have not been able to identify any in our sample.

Yi-. An example of a formal feature that affects only animate nouns is *yi-*, the initial syllable of the Dyirbal word for 'woman', *yibi*. Our sample contains

TABLE 5.3 Dyirbal nouns in *yi-*

Noun	Class	Gloss
yirrinyjila	I	'dragonfly'
yibi	II	'woman'
yigarra	II	'crayfish'
yimalimal	II	'welcome swallow'
yidir	IV	'grass'
yigan	IV	'sky'
yila	IV	'feather'
yilal	IV	'song style'
yilan	IV	'yellow feather in head of white cockatoo'
yimburr	IV	'bad smells'
yinin	IV	'wing of net trap; spirit place'
yiŋgar	IV	'long basket with cone-like mouth'
yirri	IV	'rotting material used by scrub hen to build nest'
yirribarra	IV	'nectar of forest red gum'

Dyirbal words beginning with *yi-*. As Table 5.3 shows, four nouns beginning with *yi-* in our sample are animate: *yirrinyjila* 'dragonfly', *yibi* 'woman', *yigarra* 'crayfish', and *yimalimal* 'welcome swallow'. While *yibi* and *yimalimal* are expected members of class II due to a [+female] semantic label, *yigarra* should appear in class I, the default class for non-human animates. We propose that *yigarra* was drawn into class II due to the identification of animate nouns beginning with *yi-* with class II.

The reason for the resilience of *yirrinyjila* to the formal feature is clear: the form shows vacillation of its initial segment between *y-* and *w-*. While the Dyirbal dialect of Dyirbal shows *yirrinyjila*, the Mamu dialect form is *wirrinyjila*. The same alternation is seen in one of the Dyirbal words for 'firefly', but this time between the Dyirbal dialect and Nyawagy, a closely related language; the form in the Dyirbal dialect is *yugiyam*, while its cognate in Nyawagy is *wugiyam*. Thus, it appears that the original form of both 'dragonfly' and 'firefly' began with *w-*, and that the *y-* forms are due to a development within the Dyirbal dialect.

As Table 5.3 also shows, none of the inanimate nouns beginning with *yi-* are drawn into class II; they are all in class IV, as otherwise expected. Since the source of the formal feature – the word for 'woman' – is animate, only nouns that are also animate are affected by the feature. As noted earlier, children from a very early age are able to distinguish 'animate' as a semantic category. Dyirbal shows a clear split of classification based on animacy, so it is not surprising that this split should also affect the operation of formal features.

TABLE 5.4 Dyirbal nouns in *bi-*

Noun	Class	Gloss	Motivation for assignment
bilŋgarriny	I	'little jew-fish'	animate
binyjirriny	I	'small lizard'	animate
bilmbu	I/II	'widow/widower'	male/female
bimu(nyja)	I/II	'father's elder brother/sister (and reciprocal)'	male/female
bijuju	II	'whistle duck'	female
bigi	II	'pig'	*bi-* + animate
bigin	II	'shield'	—[b]
bilmbiran	II	'large parrot'	female
bima	II	'death adder'	*bi-* + animate
binda	II	'waterfall'	water
bingay	II	'white ibis'	female
biyilbiyil	II	'peewee (magpie-lark)'	female
binana	III	'banana'	edible
binarra	III	'peanut'	edible
bigay	IV	'handle of basket'	inanimate
biguny	IV	'(finger/toe)nail'	inanimate
bilayŋgirr	IV	'blanket'	inanimate
bilbara	IV	'main track'	inanimate
bilu	IV	'noise of a horn'	inanimate
birrgil	IV	'frost, wintertime'	inanimate
biyinyji	IV	'fence'	inanimate

[b] Although *bi-* has not attracted any other inanimate nouns into class II, this formal feature could be partly responsible for the assignment of *bigin* 'shield' to class II.

Bi-. A related form for 'woman' provides another formal feature: *bi-*. A widespread word for 'woman', 'mother', and '(female) breast' in the Pama-Nyungan family is *bibi* (including Mbabaram *bib* 'breast', Muluridyi *bibi* 'breast', Nyganumarta *pipi* 'mother', Northern Nyungar *pipi* 'female breast', Kuku-Jalanji *pipi* 'breast', Kala Lagaw Ya *ipi* 'female, woman, wife'). Although the form is not in our sample, it is widespread in the region and may be related to *yibi*, discussed above. If we examine all of the forms in our sample that begin with *bi-*, we find the forms listed in Table 5.4. The number and variety of nouns beginning with *bi-* is larger than those with *yi-*, so we have also included the motivation for the class assignment of each noun. As expected, all animate nouns appear in classes I and II, and inanimate nouns appear in class III, if edible, or class IV otherwise (with the exception of *bigin* 'shield').

TABLE 5.5 Expected class I: [+male] and [+animate] nouns

Noun	Class	Gloss	Motivation for assignment
bilŋgarriny	I	'little jew-fish'	animate
binyjirriny	I	'small lizard'	animate
bilmbu	I/II	'widow/widower'	male/female
bimu(nyja)	I/II	'father's elder brother/sister (and reciprocal)'	male/female
**bigi*	II	'pig'	*bi-* + animate
**bima*[c]	II	'death adder'	*bi-* + animate

[c] While we identified *bima* 'death adder' as a potential member of the [+stinging] class, we treat it as if it is not a member of the class for the purposes of this section to show that it is doubly marked for assignment to class II.

The class assignment of most animate nouns follows straightforwardly from the semantic labels we have proposed, as shown in Tables 5.5 and 5.6. While the two starred items in Table 5.5 should be placed in class I as non-human animates, both are class II nouns. Dixon states that the assignment of *bima* 'death adder' is 'probably' due to the snake's connection to *gurrburu* 'seven sisters', a constellation also found in class II, noting that the seven sisters are 'believed to be a "death adder in the sky"' (Dixon 1972: 310). Dixon provides no explanation for the unexpected assignment of *bigi*; although *bigi* is a borrowing from English, it should nonetheless be assigned according to the principles applicable to native Dyirbal words, as shown in (7).

As Table 5.5 shows, four animate forms in *bi-* should have been placed in class I, and the actual classifications are split, with *bilŋgarriny* 'little jew-fish' and *binyjirriny* 'small lizard' in class I, while *bima* 'death adder' and *bigi* 'pig'

TABLE 5.6 Expected class II: [+female], [+water], and [+fire] nouns

Noun	Class	Gloss	Motivation for assignment
bilmbu	I/II	'widow/widower'	male/female
bimu(nyja)	I/II	'father's elder brother/sister (and reciprocal)'	male/female
bijuju	II	'whistle duck'	female
bilmbiran	II	'large parrot'	female
binda	II	'waterfall'	water
bingay	II	'white ibis'	female
biyilbiyil	II	'pee wee (magpie)'	female

appear in class II. The resilience of *bilŋgarriny* and *binyjirriny*, and the susceptibility of *bima* and *bigi*, to the formal feature is straightforward with the proper definition of the salient feature as being the composition of the entire initial syllable, not just the initial segment. The first syllable of the nouns in class II, *bima* and *bigi*, is *bi-*, while the first syllable of the nouns that were not transferred to class II is not *bi-*; both *bilŋgarriny* and *binyjirriny* begin with a closed syllable. Thus, it is the initial syllable (and not merely the first two segments) that constitutes the relevant formal feature. This is consistent with our proposal that stress (and therefore prominence) is behind the identification of certain formal features (see Section 5.4). By identifying the operation of a formal feature, we are able to explain the assignment of all four nouns without the need for tenuous semantic connections.

Ma-. An additional formal feature associated with class II is the initial segment *ma-*. Our sample in Table 5.7 contains animate nouns beginning with *ma-*. Based solely on semantic labels, *maga* 'rat' and *mawa* 'shrimp' should be assigned to class I; we have posited a [+stinging] label for *malayigarra* 'scorpion' and *marrigal* 'chicken snake', and *marraba* 'bird' is classified according to its [+female] feature.

Again, the relevant formal feature appears to be the initial syllable *ma*; the initial syllable of both forms that unexpectedly remain in class I is bigger than *ma-*, showing that it is the composition of the entire syllable, rather than its initial segments, that acts as the formal feature. However, class I contains two forms whose initial syllable is *ma-* but which resisted reclassification: *mabi* 'tree kangaroo' and *maral* 'snail-like slug'. 'Tree kangaroo' likely remained in class I due to the presence of other kangaroos, including *yunga* 'kangaroo' and

TABLE 5.7 Dyirbal animate nouns in *ma-*

Noun	Class	Gloss
mabi	I	'tree kangaroo'
mandija	I	'milky pine grub'
maral	I	'snail-like slug'
marbu	I	'louse'
maybaja	I	'alligator, crocodile'
maga	II	'rat'
malayigarra	II	'scorpion'
marraba	II	'bird (generic)'
marrigal	II	'chicken snake'
mawa	II	'shrimp'

yuri 'grey kangaroo', in class I. *Maral* similarly may have resisted reclassification due to its identification with the grubs and worms that are placed in class I. Unlike tree kangaroos and slugs, rats and shrimp may not fall as easily into a class of animate nouns identified solely with class I.

Gugu-. Another formal feature that we have identified is the initial disyllabic sequence *gugu-*. Our sample contains only three forms that begin with *gugu-*: *gugu* 'mopoke owl', *gugula* 'platypus', and *guguwuny* 'brown pigeon'. Of these three, *gugu* and *guguwuny* are birds, and accordingly placed in class II as [+female]. *Gugula*, on the other hand, has no semantic basis for assignment to class II, which Dixon was unable to explain. However, the presence of the initial disyllabic sequence *gugu-* in 'mopoke owl' and 'brown pigeon' likely was a sufficiently conspicuous feature of these class II nouns that 'platypus' was also drawn into the class.

While the majority of formal features that we have identified appear at the word's left edge, coinciding with the stressed syllable, at least one right-edge formal feature also appears to exist. We have already hypothesized that suffixes may have played a role in the original merger of small classes (see Section 5.4 above). The role of the feminine suffix *-gan* seems to continue synchronically as well. For example, the assignment of *jaŋgan* 'stonefish' to class II rather than class I, where the majority of fish species are assigned, may be due to the identification of *-gan* with class II in addition to the semantic label [+stinging], as proposed above. Table 5.8 shows all forms in our sample that end in *-gan*.

TABLE 5.8 Dyirbal forms ending in *-gan*

Noun	Class	Gloss
barrgan	I	'wallaby'
burŋgan	I	'termite species'
nuŋgan	I	'larger louse'
bugan	II	'brush fire'
julbungan	II	'woman who entices her promised man'
yalŋgayngan	II	'single woman (beyond usual marrying age)'
jaŋgan	II	'stonefish'
jarrugan	II	'scrub hen'
babuligan	IV	'pub, publican'
balgan	IV	'bark of tree'
jungan	IV	'bull oak'
bugan	IV	'open forest'
girramaygan	IV	'tribal territory'
girramaygan	I/II	'members of tribe'

With the exception of *barrgan* 'wallaby', *burŋgan* 'termite species', and *nuŋgan* 'larger louse', all animate nouns ending in *-gan* appear in class II. While the feminine suffix is clearly present in *julbungan* and *yalŋgayngan*, and likely is also present in *murrgan*, class II contains other nouns in which *-gan* would most likely be perceived but only as a formal segment, without a semantic association with [+feminine]. For example, *bugan* 'brush fire' is unlikely to contain the suffix *-gan*; it is straightforwardly assigned to class II with the other 'fire' words, and the presence of *-gan* may have been immaterial but may have also reinforced the class membership. Similarly, the assignment of *jarrugan* 'scrub hen' could be due to the belief that birds are spirits of dead women rather than the presence of *-gan*; however, this form nonetheless does end in *-gan*, implying class II membership.

If our proposal that *-gan* shifted its association from implying a [+feminine] feature to simply indicating class II membership is correct, the failure of *barrgan* 'wallaby', *burŋgan* 'termite species', and *nuŋgan* 'larger louse' to be assigned to class II deserves an explanation. As animate nouns, these should have been placed in class II due to the existence of the formal feature. First, we propose 'termite species' and 'larger louse' resisted reassignment because they form a small group of closely related insects assigned to class I, all of which are covered by a single avoidance language term, *bayi dimaniny*. Second, the different treatment of *barrgan* 'wallaby', which was not reclassified, and *jangan* 'stonefish', which was, is likely due to the relative frequencies of the forms. High-frequency forms commonly preserve irregularities that less frequent forms do not (Bybee 2002).[19]

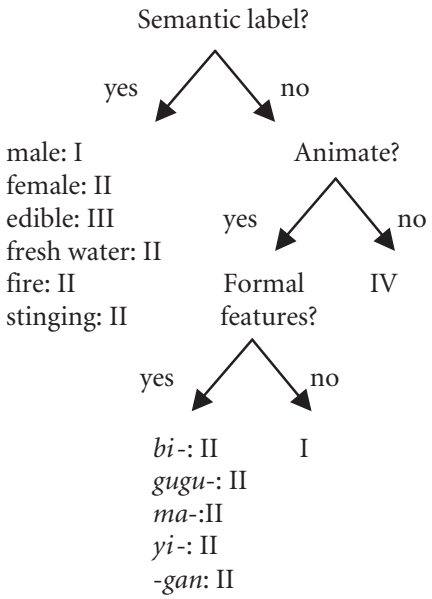
In sum, the formal features that play a role in Dyirbal gender assignment fall into two groups: stressed segments (which are word-initial), which seem to play the most prominent role, and at least one salient suffix.

To reiterate, formal features are operative only with respect to nouns that do not bear a semantic label; a noun with a semantic label is classified accordingly. However, formal features do prevent the default class assignment

[19] Although we do not have frequency numbers for the Dyirbal words in our sample, certain words would naturally be expected to be more frequent than others; for example, 'sun', 'moon', 'fire', and 'fresh water' are likely among the highest-frequency items as they were part of Dyirbal daily life. Since wallabies were hunted by Dyirbal men as sources of food, we would expect 'wallaby' to be a high-frequency item, mentioned very frequently to, and in the presence of, Dyirbal children. Although stonefish are dangerous, and accordingly Dyirbal children needed to be warned about them, they were likely mentioned with much lower frequency, only to children who had reached an age where they may encounter them, and then only as part of a warning. Thus, we propose that the failure of *barrgan* to be reclassified from the default class for animates to class II despite the presence of the formal feature we have posited is explained by the high frequency of the item.

of animate nouns that do not carry semantic labels. Accordingly, the operation of formal features in noun class assignment may be easily represented using the decision tree in (11) by adding the possibility for assignment of animate nouns based on formal features prior to assignment to the relevant default class, as shown in (12):

(12)



5.5.4 *Preservation of original class assignments*

In addition, we propose that a number of animate forms placed in class II are conservative retentions of the noun class assignments that resulted from the merger of the former Dyirbal classifiers. For example, as more fully described in Section 5.4 and Plaster and Polinsky (2007), we proposed that only 'edible' non-human animates were originally placed in class I, and that all non-edible animates fell into class II. To the extent that such items were frequent, and accordingly their class membership was conspicuous to Dyirbal speakers, such items may have resisted later reclassification to class I despite the lack of a clear semantic reason for membership in class II.

Perhaps the best example of such a noun is *guda* 'dog', whose placement in class II is unexplained by Dixon (1972). The dog held a special place in Australian culture and was not included among the 'edible' animals; as a result, it would have been excluded from the set of animals that we propose

merged with male humans to form class I.[20] Accordingly, we propose that from the beginning of Dyirbal's four-class gender system, 'dog' was a member of class II along with female humans and the other animate nouns that did not designate edible animals. While the motivation for this assignment would have been unclear to later generations of speakers after class I was reanalysed as the default class for animates rather than the class for edible animals, the frequency with which *guda* appeared in children's input ensured that it would be learned as a class II noun. Not surprisingly, once the frequency of use decreased, *guda* was reclassified as a class I noun in YD, even in the more fluent speakers of Dyirbal that Schmidt (1985) studied.

Although we have accounted for the gender assignment of the majority of the Dyirbal nouns in our sample, some unexplained forms still remain. These forms may also be conservative retentions of classifications that occurred during the merger of Dyirbal's earlier classifier classes, particularly if such items may also have been sufficiently frequent to enable them to resist reclassification – for example, the presence of *jirrga* 'eel-spear', *jumala* 'woomera', *waŋal* 'boomerang', and *warrginy* 'boomerang' in class I may indicate the existence of a class of hunting implements that merged with human males and edible animals to form class I. However, we feel there currently is not sufficient evidence to posit the existence of any such additional classes of items. We will discuss the unexplained forms in the next subsection.

5.5.5 *Outstanding forms*

We now have accounted for 573 of the 597 nouns in the Dyirbal lexicon that we had available. There is a small number of nouns outstanding (about 4%) whose gender is still unpredictable, as shown in Table 5.9. Although we do not have a definitive account of gender assignment for these nouns we would like to offer some observations.

First, the unexplained class assignments in the sample may be either conservative retentions of the original classes that resulted from the merger of the classifier system or later class reassignments, made by young language learners. However, frequency and formal features, rather than semantic content, are the likely motivators of any retention or reassignment.

Second, as Dixon (1977: 310) notes,

> it seems likely that some [class memberships] are WITHOUT EXPLANATION (as would be the case in any natural language: some may have had an explanation in

[20] Dixon (1972: 481) notes that 'people in all parts of Australia felt a close relationship to the dog (sometimes including it within the kinship system) and certainly the Yidiny would never have considered eating a dog.'

TABLE 5.9 Unexplained classifications

Noun	Class	Gloss
burrubay	I	'boil'
gubaguba	I	'(type of stripy pearl shell)'
jirrga	I	'eel-spear'
jumala	I	'woomera'
maŋany	I	'boomerang'
mayjala	I	'flash of lightning'
mayjanmayjan	I	'continuous flicks of lightning'
mugay	I	'grinding stone'
waŋal	I	'boomerang'
warrginy	I	'boomerang'
balma	II	'old scrub hen nest'
baŋgay	II	'spear'
buluba	II	'fighting ground'
bulugi	II	'cattle'
bunarra	II	'bow and arrow'
bundiny	II	'grasshopper'
gabu	II	'cup, telephone/telegraph mouthpiece or earpiece'
galabay	II	'beetle'
gawa	II	'cow'
jayari	II	'horse'
lambi	II	'lamp'
nyiyi	II	'noise of birds'
ŋama	II	'shield handle, trigger (on gun)'
warrayi	II	'bony bream'

terms of an earlier stage of the language, but the class assignment has been retained and the explanation lost as the language has altered).

While we have been able to reach earlier stages of Dyirbal to explain certain class assignments that were not predicted by Dixon's classification system and to motivate certain 'myth-and-belief' assignments on a more solid basis, we have not been able to determine the explanation for all items. Nonetheless, our successes demonstrate that the inquiry is worthwhile, and may shed light not only on the linguistic development of the language but also the source of certain pieces of myth and folklore.

Third, there may be additional semantically motivated classes that are difficult to ascertain because of the small size of the sample. For instance, it

is feasible that cows, horses, and pigs all formed a coherent small class of domesticated animals, which would explain the assignment of both *gawa* and *jayari* to class II, where *guda* 'dog', the other domesticated animal, is found, and which could have served as an attractor. At the current stage of our knowledge of Dyirbal vocabulary, such proposals are doomed to be speculative. We prefer to leave some items unexplained rather than posit additional classes on the basis of limited data.

5.6 What happens when categorization collapses: noun classes in Young People's Dyirbal

As mentioned above, Schmidt (1985) documented the language of the children and grandchildren of Dixon's Dyirbal informants and identified two groups of speakers: a more fluent group of speakers of traditional Dyirbal, and a less proficient group of speakers of a new variety of the language, Young People's Dyirbal. While the more fluent speakers preserve a majority of the features of traditional Dyirbal, YD has undergone a variety of changes that differentiate it from traditional Dyirbal. Most relevant for this chapter is the simplification of the noun class system into a straightforward three-class system, as set forth in (13). The reanalysis of Dyirbal noun classification was largely driven by the concomitant process of attrition in noun class agreement – with diminishing agreement, speakers had fewer cues that allowed them to differentiate between several noun classes. Such a reanalysis with further simplification based on salient linguistic characteristics is attested in different languages undergoing attrition, for example, in Gaelic (Dorian 1980) and Heritage Russian (Polinsky 2008).

(13) Noun classification in YD (Schmidt 1985: 158)

Category		Class
animate:	masculine	I
	feminine	II
inanimate		IV

Schmidt found that YD speakers classify nouns solely on the basis of the categories set forth in (13), and, as a result, reclassify all nouns whose former classification did not comply with the YD system. For example, YD speakers retain none of the mythological class assignments found in traditional Dyirbal, as shown in Table 5.10.

TABLE 5.10 Noun classification by traditional Dyirbal and Young People's Dyirbal speakers (Schmidt 1985)

	Traditional Dyirbal	Young People's Dyirbal
'man'	I	I
'rainbow'	I	IV
'moon'	I	IV
'storm'	I	IV
'woman'	II	II
'bird'	II	I
'sun'	II	IV
'star'	II	IV

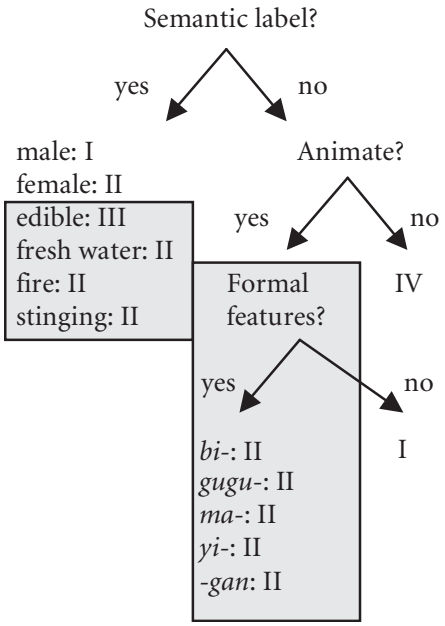
As Table 5.10 shows, in YD all inanimate nouns appear in class IV, despite any semantic associations that caused them to be placed in class I or II in traditional Dyirbal. Similarly, birds appear in class I with all other non-human animate beings. These developments, which occurred over a relatively short period of time, underscore the fragility of the hypothetical mythological associations in noun categorization.

In addition, YD speakers place the formerly [+fresh water] and [+fire] items in class IV rather than class II. This reassignment is expected under our analysis, due to the loss of the more minor semantic labels.

For prior Dyirbal speakers these labels were both relevant and frequent, and as a result children quickly identified the semantic classes and their class assignment. YD speakers are much less likely to hear or participate in conversations about fires than their ancestors, who cooked on fires and slept near fires for warmth. Sources of fresh water are likely no longer frequently mentioned, and water may not be the first beverage that comes to their mind when they are thirsty. Schmidt's YD speakers simply did not receive sufficient input to motivate the extra semantic labels on these inanimate items; as a result, they are placed in the default class for inanimate nouns.

Returning to the system that we proposed for noun classification in traditional Dyirbal in (12), we find that YD has preserved the core of the basic categorization that was present in the traditional system, while eliminating the formal features and removing the more ancillary semantic labels that were relevant to traditional Dyirbal speakers. These abandoned classification devices appear in shaded boxes in the decision tree set forth in (14); the remaining decision tree represents the classification system of the YD speakers:

(14)



Thus, the development of the Young People's Dyirbal noun class system follows directly from the noun class system of traditional Dyirbal, with the loss of the more ancillary classification devices. The changes seen in the YD noun class system need not be attributed to the direct interference of English or the discontinuation of the traditional Dyirbal worldview, as has been suggested; rather, while the abandoned pieces of the tree were able to be maintained by Dixon's informants, their children and grandchildren did not receive sufficient input to infer the existence of these semantic categories and to identify these formal features.

5.7 Conclusion

In this chapter, we have reanalysed Dyirbal noun classification, which has been previously analysed in terms of radial categories that rely on complex conceptual associations. We have proposed that the assignment of Dyirbal nouns to genders is determined by a combination of rather straightforward semantic and formal features. Crucially, the relevant semantic features are quite similar to what is found in gender systems around the globe; they include the basic semantic labels 'animate', 'male', 'female', and 'edible'. These features comprise the semantic core of Dyirbal classes I through III.

The fourth class, which includes most of the inanimate nouns in the lexicon, constitutes the default gender for nouns that do not bear these labels.

We have also proposed a small subset of minor semantic labels, such as 'water' or 'fire', which identify semantic subclasses within larger classes. The Dyirbal noun classes arose from a reanalysis of an earlier classifier system; the original number of classifiers was larger than the number of resulting genders, and in several cases, several classifier sets merged within a single class. We hypothesize that this merger was facilitated by formal analogy between the members of different small classes. If this proposal is on the right track, it has as an important consequence that there is no synchronic conceptual association among all of the items in a given gender class; in particular, the smaller subsets within a class do not need to be radially related to the semantic core. The overall class membership is motivated only diachronically, and even then not necessarily on semantic grounds.

We have identified several formal cues that play a role in synchronic gender assignment in Dyirbal and may have also affected the merger of smaller classes in the history of the language. Most of these formal cues are provided by the initial syllable, which is where the primary stress falls in Dyirbal; we have been able to connect at least some of these salient formal cues with words that represent core semantic notions. The reliance on prosodically prominent word segments is known to motivate gender assignment in many unrelated languages, including familiar Indo-European languages like French and Latin, so this finding also brings Dyirbal closer to well-known (and rather unsurprising) gender systems. Of course, any reconstruction remains tentative and hypothetical, and, in a linguistically complex and documentally impoverished area such as Australia, the tentative nature of any proposal needs to be underscored; we certainly leave room for doubt in our proposal, but the existing body of data and various cross-linguistic parallels make it plausible.

The development of the straightforward gender system of Young People's Dyirbal directly follows from the explanation of traditional Dyirbal gender proposed herein. While noun classification in traditional Dyirbal involved several semantic and formal features, a noun's gender is determined in YD solely on the basis of the objective, readily apparent characteristics of its referent: masculine animate beings are placed in class I; feminine animate beings are placed in class II; and everything else falls into class IV. The changes between traditional Dyirbal and YD are precisely those that we would expect given the decreasing use of the language in the younger generation.

From the standpoint of learnability, the proposed account of Dyirbal genders is more plausible than one based on attenuated abstract semantic linking. Children show early acquisition of superordinate categories but are

less likely to acquire more sophisticated and culture-specific semantic categorization at an early age (Mandler 2004). Children are also known to pay attention to statistical and phonetic cues in their language in the first year of life (Jusczyk et al. 1993, 1994; Saffran et al. 1996a, b; Karmiloff-Smith 1979; Levy 1983; Berman 1985; Smoczyńska 1985; Slobin 1973; Newport and Aslin 2000, 2004). Although adult speakers may offer intriguing generalizations concerning the motivations for gender assignment, there is no evidence that these are any more than after-the-fact rationalizations that speakers of any language often come up with, let alone what children use to assign gender.

Despite its initial appearance, Dyirbal gender does not require complex semantic rules and links; instead, it relies on core semantic categories and independently motivated features such as stressed syllables and salient suffixes, and is directly comparable to other noun classification systems, which often rely on a combination of formal and simple semantic cues (Corbett 1991: Chapters 2 and 3). Even though the language is not yet officially extinct, working with Dyirbal is much like working with a meagrely attested ancient language: we are extremely thankful for what has been preserved (and to Dixon for his many efforts to do so), and we must extract everything possible from the surviving attestations, even if they cannot provide all of the answers we seek. We cannot help but wonder what would have become of the Dyirbal gender system without the drastic impact of the colonization of Australia on the language and its speakers.