

## 2

---

# Unnatural kinds

MARK ARONOFF

### 2.1 Natural kinds and natural language

Are the categories of languages, especially morphological categories, natural in any sense of the term? The term *natural kind* comes from philosophy and I will begin to try to answer my question by examining the term *natural kind* as it is used in that field. From there, I will move to a historical discussion of the term *natural language*. Within modern linguistics, there is a tradition according to which the categories of natural languages are said to be natural, in the sense of being selected from a predetermined set given by (human) nature. I will explore the success of this idea within phonology and morphosyntax and then move on to morphological categories, arguing that it is not useful to regard all morphological categories as natural. I will close with a discussion of morphological and lexical categories in sign languages, showing how these categories are influenced by human interaction with the environment. It has been suggested to me that perhaps some but not all of the categories of natural languages are preordained, selected from a predetermined set. Such a position is a bit like saying that only some species were created and others derived. If some of the categories of natural languages are demonstrably unnatural, in the sense explored in this article, then it is most reasonable to assume that none of them are. The question then arises of how categories emerge if they are not given in advance. I will not address this question here.

#### 2.1.1 *Natural kinds*

The term *natural kind* comes from philosophy, more specifically ontology, the branch of philosophy that is concerned with what truly exists. A natural kind is a grouping that is not created by human culture or imagination but that exists independently. According to scientific realism, when a science is successful, the kinds of entities posited by that science will correspond to the real kinds in nature and will hence be natural.

Species, for example, used to be considered immutable kinds given for all time, but Darwinian evolutionary biology showed that this was not the case, the most dramatic demonstration that it is not at all obvious what kinds are natural. A kind does not have to be simple in order to be natural. Water, for example, is not simple, since it consists of a combination of hydrogen and oxygen, but it is still a natural kind, since it is not created by human culture and there are good scientific reasons to believe that it exists independently of humans.

We can trace a progression in the discussion of natural kinds from metaphysics to linguistics through philosophical semantics. Although the debate can be traced back at least as far as medieval philosophers like Duns Scotus and William of Ockham, it was also a flourishing concern among early modern metaphysicians. John Stuart Mill, for example (Mill 1884), asked whether the kinds that we think of as ‘natural’ kinds are genuinely *natural*. In the twentieth century, Saul Kripke (1980) and Hilary Putnam (1973, 1975) sought to construct a semantic theory of reference for natural kind terms that does not depend on internal mental representations. Cognitive psychologists and linguists turned this question on its head, asking not about kinds in nature but rather about kinds in cognition: whether there are certain kinds or categories that are cognitively more likely or that are more natural for children to learn and language to express (Mervis and Rosch 1981). For linguists, this can be thought of as asking whether certain linguistic categories are natural to humans, part of our nature.

### 2.1.2 *Natural languages*

Whether there is any connection between the kinds, if any, that are natural to humans and the kinds that exist in nature is a separate question that we will not address, but the whole notion of a natural language has its origins in this very issue. In 1668, John Wilkins, one of the founders of the Royal Society, published *An essay towards a real character and a philosophical language*. Wilkins proposed what he called the Real Character as a universal set of ideographic visual symbols vaguely resembling Chinese but not meant to be pronounced, each representing a basic thing or notion. It was real in the sense that the symbols purported to correspond to reality. The Philosophical Language assigned phonetic value to each of the real characters. Wilkins’s system is of interest to the history of science in so far as it was a precursor to Linnaeus. For linguists, the most interesting aspect of the work is that it raises the possibility of a universal natural language.

Wilkins was influential enough to be satirized in Jonathan Swift’s *Gulliver’s Travels* a half century later in 1726. The relevant passage is from Chapter V of Part III: A Voyage to Laputa, Balnibarbi, Luggnagg, Glubbdubdrib, and Japan:

The other project was, a scheme for entirely abolishing all words whatsoever; and this was urged as a great advantage in point of health, as well as brevity. For it is plain, that every word

we speak is, in some degree, a diminution of our lunge by corrosion, and, consequently, contributes to the shortening of our lives. An expedient was therefore offered, ‘that since words are only names for things, it would be more convenient for all men to carry about them such things as were necessary to express a particular business they are to discourse on’.

... many of the most learned and wise adhere to the new scheme of expressing themselves by things; which has only this inconvenience attending it, that if a man’s business be very great, and of various kinds, he must be obliged, in proportion, to carry a greater bundle of things upon his back, unless he can afford one or two strong servants to attend him. I have often beheld two of those sages almost sinking under the weight of their packs, like pedlars among us, who, when they met in the street, would lay down their loads, open their sacks, and hold conversation for an hour together; then put up their implements, help each other to resume their burdens, and take their leave.

Nineteenth-century grammarians from Schleicher to Baudouin de Courtenay stressed the notion that linguistics is a natural science because its object, human language, is natural. Saussure (1959) placed great emphasis on the primacy of spoken language, in large part because of the historically contingent nature of writing:

The concrete object of linguistic science is the social product deposited in the brain of each individual, i.e. language. . . . The linguistic object is not both the written and the spoken forms of words; the spoken forms alone constitute the object. (Saussure 1959: 23–4)

Forty years later, Chomsky in *The logical structure of linguistic theory*, written in the mid-1950s, though not published until twenty years later, used the term *natural language* to name the object of linguistic inquiry, referring to ‘all “humanly possible languages”—in particular the attested natural languages’ (1975: 13), identifying the attested natural languages as a subset of those which are humanly possible, which he took to be the real object of inquiry.

If humanly possible languages are natural, are they made up of natural kinds of objects? One type of positive answer posits that the set of possible natural languages is defined by a feature or category space, where the features or categories are natural (to humans) and innate. This is more or less the principles and parameters view of human language (Chomsky 1981, Baker 2002). There are a number of good reasons for this sort of answer. The first is **universality**: positing innate features is one simple way to encode the observation that all natural human languages share certain fundamental properties: they share them because they are innate. The second is accounting for **variation**: natural human languages vary from one another but positing innate universal features may allow us to constrain the universe of variation. A third is **ease of comparison**: if there is a fixed universal of features in a given domain, it is easier to compare natural human languages to one another than if languages can vary freely. A fourth is **learnability**: if something is innate, we don’t have to learn it, at some level of analysis. There may be other reasons, but these have been most emphasized.

### 2.1.3 *Natural syntax*

One linguistic theory that has made especial use of innate features is that of cartographic syntax (Cinque and Rizzi 2010), which is concerned with functional categories, especially those involved in verb inflection. Shlonsky (2010) writes that ‘familiar arguments from the poverty of stimulus militate in favour of the hypothesis that the functional features and their hierarchical arrangement are wired into the grammar and not learned or otherwise acquired through experience’.

Shlonsky makes his reasoning clearer in the following passage:

The question of what is coded or articulated by grammatical features can be rephrased as following: Of the properties which enter into human thought and belief systems, which ones are represented as grammatical features? Although we are quite far from seriously addressing this question from the perspective of studies of human perception/cognition, comparative and typological studies of particular grammars can provide a rough estimate of the magnitude of the inventory of the features involved as well as of their degree of granularity. Kayne (2005b: 12–14) lists about 40 but notes that that the estimate is conservative in that it fails to take into consideration features which are unpronounced in languages he has come across and that, moreover, some mono-morphemic forms arguably contain more than a single feature. Citing Heine and Kuteva’s (2002) study of grammaticalization targets, Cinque and Rizzi (2008) raise the order of magnitude by 10, to about 400. (Shlonsky 2010: 424)

This line of reasoning, according to which all the categories that are natural to humans must be given in advance, inevitably leads to the multiplication of entities, unless we assume that the categories are themselves made up of smaller atoms, which is precisely the tack that has always been taken in phonological feature theory, which I will now turn to.

### 2.1.4 *Natural phonology and binary phonological features*

Modern phonological theory has its origins in Saussure’s notion of opposition and value (Saussure 1959): ‘in classifying phonemes, what constitutes them is of much less importance than what distinguishes them from each other’ (1959: 49). Trubetzkoy coined the term *distinctive opposition* in the 1930s. I quote here from Cantineau’s (1949) French translation: ‘Les oppositions phoniques qui dans la langue en question peuvent différencier les significations intellectuelles de deux mots, nous les nommons des *oppositions phonologiques*<sup>1</sup> (ou des *oppositions phonologiques distinctives* ou encore des *oppositions distinctives*’ (Trubetzkoy 1949: 33). [The sound oppositions that in the language in question can differentiate the mental signification of two words, we will call *phonological oppositions* (or *distinctive phonological oppositions* or even *distinctive oppositions*]. In the footnote to this passage, he suggests English terminology: ‘Pour l’anglais par contre nous recommanderions l’expression “distinctive opposition”’,

car aussi bien “phonological opposition” que “phonemical opposition” peuvent être mal compris.’ [For English on the other hand we would recommend the expression ‘distinctive opposition’, because ‘phonological opposition’ as well as ‘phonemical opposition’ can be misunderstood.] Jakobson extended Trubetzkoy’s programme, suggesting that phonemes should be decomposed, a move that is in line with Jakobson’s general atomistic line of thought and his search for invariance in language: ‘Nous identifions les phonèmes d’une langue donnée en les décomposant en leurs caractères phonologiques constitutifs, c’est-à-dire que nous établissons pour chaque phonème quelles *qualités* l’opposent aux autres phonèmes du système en question.’ [We identify the phonemes of a given language by decomposing them into their constitutive characteristics, i.e. we establish for each phoneme those *qualities* that oppose it to all other phonemes in the system under consideration] (Jakobson 1962: 272). In this same article, with an explicit nod to Trubetzkoy, he introduces the idea of decoding all phonological differences into binary oppositions.

A decade later, Jakobson, Fant, and Halle set out their hypothesis: ‘The inherent distinctive features which we detect in the languages of the world and which underlie their entire lexical and morphological stock amount to twelve binary oppositions’ (Jakobson, Fant, and Halle 1952: 40). Halle reiterated this position in the founding work of generative phonology: ‘the phonetic properties in terms of which segments are characterized belong to a specific, narrowly restricted set of such properties called the *distinctive features*. All distinctive features are binary’ (Halle 1959: 19). The idea that these features are exhaustive also remains a hallmark of the theory: ‘the totality of phonetic features can be said to represent the speech-producing capabilities of the human vocal apparatus’ (Chomsky and Halle 1968: 297).

In many ways, Chomsky and Halle (1968) was the high-water mark of universal binary-feature-based generative phonology. Both binarity and universality have since been questioned. Peter Ladefoged, for example, long questioned the notion that all phonetic features must be binary, as exemplified in the following passage from the last piece he published: ‘The phonetic input can be defined in terms of the values of twenty-two articulatory features, each of which is a specific phonetic parameter. Languages can use different values along these parameters to distinguish words, but there are fifty-five named values representing modal possibilities’ (Ladefoged 2007: 178). Ladefoged notes that nine of his twenty-two features have more than two values and he lists five values for the feature of vocalic height and glottal stricture. But even these values are not absolute, but modal: ‘There is a continuous range of values within a feature’ (Ladefoged 2007: 168). Freed from binarity (which was always rooted in method rather than in fact), the next logical bond to break is that of universality/innateness. Mielke (2005) is a good example of this move: ‘natural classes emerge in a model that has access to the observable phonetic properties of sounds, but no innate

features. Innate features are not needed to rule out phonetically unnatural classes' (Mielke 2005: 289). Pierrehumbert (2003) is similar in spirit, though not so explicit in the discussion of natural classes.

It is important to emphasize that what Mielke calls natural classes are not natural kinds in any normal sense of the term. They are categories that are grounded in human nature and experience, but they are not ontologically natural or given in advance. Instead, they emerge from an interaction between the learner and the input. The unfortunate continued use of the term *natural class* in phonology into the present time has led to a great deal of confusion about the basis of the entire enterprise.

Looking back over more than a half century of phonological research based on a fixed universal set of binary distinctive features, it is clear that the search for innate universal features that characterize all languages has not been as successful as originally hoped and may have been rooted in an overly simple view of the limits of human cognition. Is there some sense of natural that might be helpful in the study of languages that does not depend on the innateness of categories?

## 2.2 Embodied categories

One important trend in cognitive research over the last quarter century is the study of embodiment. The term *embodiment* has many definitions and ranges across a wide variety of fields, from phenomenology (Rowlands 2010) to law (Fletcher, Fox, and McCandless 2008). In this work, I will assume the least controversial of these definitions, taken from Margaret Wilson's now classic 2002 article: 'The mind must be understood in the context of its relationship to a physical body that interacts with the world' (Wilson 2002: 625). When we say that language is embodied in this conservative sense we are claiming only that, because natural languages are produced by the human mind/body, they are constrained by both human nature and how our minds/bodies exist in and interact with the world. Human languages are the way that they are at least in part because they are produced by human minds/bodies situated in the world that we inhabit.

How do we cash out this position in a discussion of the naturalness or unnaturalness of linguistic categories? First, if the categories of natural languages are rooted in human nature, then this nature and these categories can be studied by investigating the human mind/body and its relation to the world. Second, we don't have to assume that the categories themselves are natural kinds in the strong sense of being determined in advance, only that they are compatible with and in some cases determined by how the human mind/body interacts with the world. This view of what is natural about language may lack the explanatory advantages that positing innate categories and values gives (discussed above), but it may provide a better fit to what we actually find in human languages and it avoids the methodological pitfalls inherent in positing a large number of entities.

2.2.1 *P(erson), N(umber), and G(ender)*

As a test case, I will consider the P(erson), N(umber), and G(ender) features that participate in verb agreement, often called  $\Phi$ -features in the formal syntax literature. I will show, in an abbreviated fashion, that these features form a cline of naturalness, with Person invariant and firmly rooted in discourse, Number more variable but still clearly adhering to a universal cognitive system, and finally Gender based in some sense on cognitive principles but in no sense determinate at all.

Consider, for starters, the variety of values of the feature person across languages. Siewierska (2004), based on a survey of over 700 languages, concludes that there are universally three values for person: first, second, and third. She characterizes first and second person as expressing ‘the participant or discourse roles of speaker and addressee’ (Siewierska 2004: 1) and the last as a third party in a discourse. First and second person pronouns are inherently deictic, while third person pronouns are more often used anaphorically rather than deictically. Further distinctions within third persons are sometimes made and the term *fourth person* is sometimes used for some of these but Siewierska concludes that ‘Under none of the above uses does the fourth person qualify as a bonafide additional discourse category’ (Siewierska 2004: 7). In other words, the three universal values of the grammatical feature person are rooted in the nature of discourse or conversation. Wechsler (2010) shows that this sort of treatment, in particular one in which first and second person pronouns, unlike other nominals, are inherently deictic or indexical, accounts for the peculiar semantic properties of first person plurals, which is that they always allow for the inclusion of a referent other than the speaker or hearer. Wechsler shows that if we instead posit innate universal feature values like 1 and 2 (and maybe 3), which are not derived from the communicative context, what he calls a *UG solution*, we fail to express this special associative property of first and second person pronouns in all languages. In this case, the nature of conversational interaction provides an explanation for the person values that are actually found in natural languages.

Grammatical number values, while more diverse than person, are also quite limited in their variety. Corbett (2000) shows that a language may show either no grammatical number distinction, a distinction between singular and plural (as in English), a further distinction of dual (Classical Greek), a further distinction of trial or paucal beyond dual 1, or it may distinguish both trial and paucal (Lihir) beyond dual. More formally:  $(\text{trial} \vee \text{paucal}) \supset \text{dual} \supset (\text{singular} \wedge \text{plural})$ . This typology accounts for all the grammatical number systems in Corbett’s sample of some 250 languages. It is consistent with the findings of Dehaene et al. (1999) and Frank et al. (2008) that, while humans have a universal ability to estimate approximate quantity, exact counting systems are culturally determined, though they all follow the basic Peano axioms for the natural numbers. Grammatical number systems, then, while not as limited in their variety as person systems, are still grounded in both human

cognition and the laws of mathematics. In order to talk about grammatical number, we still need no special features that are particular to language.

We come now to grammatical gender. Here, a quote from a classic article on phonology is apt:

Thus, languages have rules which are plausible or which can be derived from plausible rules by a sequence of steps... but in the process rules can become highly implausible. In short, languages have crazy rules. (Bach and Harms 1972: 6)

Genders, it turns out, are like crazy rules: they happen. They start out grounded in plausible cognitive categories but most gender systems eventually shuffle off this mortal coil. Corbett notes near the beginning of his book *Gender* that:

In a sense all gender systems are semantic in that there is always a semantic core to the assignment system (Aksenov 1984: 17–18). (Corbett 1991: 8)

Corbett is referring here to an article whose title is *K probleme èkstralingvistièeskoj motivacii grammatičeskoj kategorii roda* [On the problem of the extralinguistic motivation of grammatical gender]. From this title it is clear that by *semantic* Corbett is referring to extralinguistic motivation. This, as I will show, is the only sense in which genders could be natural. If instead we try to claim that genders are natural in the sense of being selected from a universal set of innate feature values, we quickly come to grief.

All gender systems may be traced back ultimately to some semantic/cognitive categories, but there are, according to Corbett, very few strict semantic systems, in which the meaning of a noun is necessary and sufficient to determine its gender, even though this is precisely the sort of ‘natural’ gender system that one might naively expect to be most common. The Dravidian languages and some North-East Caucasian languages fall into this type, distinguishing the cognitive categories male rational, female rational, and non-rational as three genders. In some Dravidian languages, e.g. Telugu, the female rational and non-rational have fallen together. Corbett mentions a few other strict semantic systems, including English, all of which obey this male, female, non-rational distinction.<sup>1</sup> Next come ‘predominantly semantic systems’, in which one can often predict the gender of a noun from its meaning by ‘gender assignment rules’, but the rules do not cover all nouns. Corbett gives the example of Zande (Niger-Congo), which has four genders. Male human and female human are completely predictable both ways but the animate non-human gender, though including all animate non-humans, also includes some inanimates, which are for the most part individual idiosyncratic exceptions (the remaining inanimates form

<sup>1</sup> For English, at least, this system only works for third person pronouns, making it a very limited gender system.

the fourth gender). This sort of pattern, with two completely predictable genders and two less so, turns out to be fairly common. The actual distinctions, Corbett notes, fall to some extent into a series of dichotomies first enumerated by de la Grasserie (1898) based on humanness, animacy, gender, and relative size or power. But Corbett notes that this schema is far from exhaustive and lists examples from different languages of genders for insects, canines, hunting weapons, and liquids. Thus, while there are languages whose gender systems are strictly semantic, they are few in number and the set of semantic genders is not as small as one might wish.

Worse, though, is the fact that many gender systems, although they have semantic assignment principles, also have what Corbett calls ‘formal assignment rules, that is, rules which depend on the form of the nouns involved rather than their meaning’ (33), i.e. morphology and phonology. Phonological rules depend on the sound-form of the word, while morphological rules may depend on inflectional class or particular affixes; for example, the German suffix *-chen* assigns neuter gender, so that words like *Mädchen* ‘girl’ or *Männchen* ‘little man’ are neuter in gender, even though they refer to a human female and male, the semantic archetype for the feminine and masculine genders. Furthermore phonological gender assignment is sometimes statistical in nature, as demonstrated in detail for French by Tucker, Lambert, and Rigault (1977).

There are other types of assignment rules. In Classical Greek, indeclinable nouns, like the letters of the alphabet, whose names are borrowed directly from Phoenician, are neuter. In Russian, indeclinables are masculine if they designate animates, but neuter otherwise. The assignment of loanwords to genders follows complex patterns in many languages, some of which are discussed in detail by Corbett. Often, loanwords pattern according to statistical tendencies already present in the language. For example, in German, monosyllabic borrowings are masculine, presumably because the majority of monosyllables are masculine (Carstensen 1980), resulting in a minimal pair of sorts between *das Lied* and *der Song*, which are almost synonyms.<sup>2</sup> In languages with alliterative genders, a loanword with a phonological form that does not fall into any of the genders of the language will usually be put into the default gender (Aronoff 1994) but sometimes the agreeing element will show concord simply by alliterating the initial CV of the noun (Sauvageot 1967, 1987; Dobrin 1995, 1998).

### 2.2.2 *Gender assignment and semantics*

In the end, what matters most for gender systems is not what their basis is, but that the genders be exhaustive: every noun must be assigned a gender so that every agreeing element can have its proper form. How the genders are assigned may be quite complex, though. Corbett examines two gender systems in detail, those of Russian and Swahili.

<sup>2</sup> Thanks to Mark Lindsay for this example.

In each system, semantic gender assignment rules are accompanied by morphological rules based on inflectional class, including, for Russian, whether a noun can be inflected or not. And semantics does not always take precedent, as we saw above for German, where individual suffixes override semantics.

Some gender systems abound in minor semantic gender assignment rules unrelated to the ‘core’ semantics of the gender system. Consider Latin, which has three genders, masculine, feminine, and neuter:

- Latin names of rivers (*Se:quana*, *Tamesa*) and winds (*Aquilo*, *Corus*) are masculine (*flumen*, the word for ‘river’, is neuter; *ventus*, the word for ‘wind’, is masculine).
- Latin names of trees (*fagus/i*, *laurus/i*, *pinus/i/u:s*, *quercus/u:s*) are feminine (*arbor*, the word for ‘tree’, is feminine).
- Latin names of fruits are neuter but names of fruit trees are feminine (*ma:lum* vs *ma:lus*; *pirum* vs *pirus*; *amygdalum/a* vs *amygdala*), following the general pattern for tree names.

Examples like this help us to understand why the Latin names for the ‘male’ and ‘female’ genders are diminutives: *masc-ul-in-us* (double diminutive!) and *femin-in-us*. The diminutive here has the force of the English *-ish* suffix. The gender names mean ‘maleish’ and ‘femaleish’, because, even when gender assignment is semantic, many of the items that fall within each of these genders do not denote male or female entities. The word *neuter* by contrast means simply ‘neither’, which is what it is.

In Aronoff (1994), I discussed in detail the assignment rules for Arapesh, a language with many genders and inflectional classes, in which semantic rules assign the human genders and the human genders assign inflectional classes, but where for most nouns the assignment goes in the opposite direction: from phonology to inflectional class to gender. Again, what matters most is that the gender system be exhaustive, that each noun be given a gender, not whether there is any coherent system for assigning nouns to genders.

### 2.2.3 Using PNG

It is important to distinguish the semantics of PNG categories of a language from how they are used.<sup>3</sup> In contemporary English, for example, the most impersonal form is *you*. This use of *you* is not third person (*you don't/\*doesn't do that sort of thing around here*). In more formal discourse, *we* can be used as well but is not third person either (*If we assume/\*assumes that all swans are white*). There is an impersonal pronoun (*one*) but it is used only in the most stilted of styles.

<sup>3</sup> Thanks to Irit Meir for reminding me and for referring me to the Israeli Hebrew examples.

In many languages, there is a special impersonal pronoun, like German *man* (not to be confused with the etymologically identical *Mann*) or French *on*. Other languages have no purely impersonal pronoun and the impersonal is expressed by some other third person pronoun, singular or plural. The major Romance languages other than French use the etymological reflexive, *si* in Italian, *se* in Spanish. In English, third person plural *they* can be used in expressions like *they say*. The English usage contrasts nicely with Russian or Israeli Hebrew, in which the impersonal is expressed by a third person plural verb, but with a null subject. The Hebrew equivalent of impersonal *they say* is *ʔomrim* ‘say.present.plural.masculine’; with the overt subject (*hem ʔomrim*) only a referential sense is possible. Nor can the feminine plural be used.

In French, by contrast, the impersonal pronoun *on* can be used ‘instead’ of the first person plural, though often accompanied by the first person plural pronoun: (*nous on va à New York* ‘we are going to New York’ but literally ‘we one is going to New York’). It is not clear what semantic difference if any there is between this sentence and the equivalent *Nous allons à New York*.

Polite second person expressions are notoriously peculiar. In Italian, the third person feminine pronoun *lei* is always used for polite second person singular reference; it triggers third person agreement but not feminine agreement, unless the referent of the subject is female: *lei è simpatico* ‘you m. are friendly’ but *lei è simpatica* ‘you f. are friendly’.

A nice example of the conventionalized use of gender comes from Israeli Hebrew, where masculine and feminine are obligatorily distinguished in the second person singular form of the verb and even in the first person singular (though only in the present tense). Tobin (2001) provides a detailed exposition of the use of the masculine form as a device to indicate affection or intimacy in addressing a female. His examples range from a conversation between lovers in a novel to transcriptions of a number of interactions between eight-year-old twin sisters. To someone unfamiliar with the usage, the passages are extremely jarring.

All of these uses are conventional. English speakers may believe that their use of the second person in impersonal contexts is perfectly natural; German or French speakers regard it as peculiar. Israeli Hebrew speakers think that there is something normal about the gender switch and are surprised that it does not occur in other languages.

#### 2.2.4 PNG and natural kinds

Which brings us back to natural kinds. It should be obvious from this discussion of gender assignment that theorizing about genders in terms of some universal set is just not helpful. Even when two languages ostensibly have the same genders, unless the genders are strictly semantic (which is rare), the languages will differ in how

nouns are assigned to them. Latin and German, for example, both have the same three original Indo-European genders (masculine, feminine, and neuter), but aside from the default semantic assignment of human males and females to masculine and feminine, the two systems are quite different. It is also very hard to see how to fit phonologically driven genders of the sort found in Arapesh and Yimas (Aronoff 1994) into any meaningful universal scheme. The considerations that might lead one to posit a set of innate humanly natural categories (universality, variation, ease of comparison, and learnability) lead us here to the opposite position: that gender, as opposed to person and number, when it occurs, is by and large a language-particular category. It is noteworthy that, while, according to the *World Atlas of Language Structures* all languages have person systems, ten per cent (28 languages out of a sample of 291) have no nominal number marking, and the majority of languages (145 out of a sample of 257) have no grammatical gender distinction at all. Since grammatical gender is found in less than 45 per cent of a representative sample of the world's languages, we should not be surprised that when it does occur, it is so variable.

If we think only about the two types of assignment rules, morphological and phonological, that Corbett calls 'formal' because they depend on the form of a noun, rather than its meaning, such rules, and the gender categories that they map onto, are by definition not natural, because they depend on the arbitrary properties of the forms of words. Even worse are the 'formal' rules that assign gender on the basis of inflectional class, since they don't even have anything to do with real form but only with whether a word belongs to a particular arbitrary class of words that share a certain inflectional paradigm. Let us refer to these gender categories as unnatural, as opposed to natural. We now see that languages have unnatural kinds. Why languages should have unnatural kinds can best be understood in terms of the central human attribute, culture.

### 2.3 Culture and unnatural acts

Linguistics in the last half century has de-emphasized the connection between the two great human attributes, language and culture, in large part because of a desire to emphasize the extent to which languages resemble one another, as opposed to cultures, which differ so greatly. E. B. Tylor, the founder of cultural anthropology, famously defined culture as follows: 'Culture, or civilization, taken in its broad, ethnographic sense, is that complex whole which includes knowledge, belief, art, morals, law, custom, and any other capabilities and habits acquired by man as a member of society' (1871: 1). Put more succinctly, we may define culture as those capabilities and habits that are determined by society rather than by nature. There is evidence for rudimentary culture in other animals (Laland and Galef 2009), but if we ask what is fundamental to human nature, then what sets us apart from all other

creatures is the extent to which what we do goes beyond nature and to which we codify and explicitly value these unnatural doings. Sterelny (2009) calls this property, which sets humans apart from all other animals, even animals for which we have evidence of socially transmitted behavioral traditions, *enculturation*. It is in our nature to be encultured, to go beyond what nature gives us and to adopt norms of behaviour and belief. Hill (2009) presents a similar point of view. We may accordingly define culture as consisting of unnatural acts, systems, and beliefs.

Culture, by Tylor's definition, includes language and, just like any other aspect of cultures, languages depend on human nature but are not determined by it. The modern emphasis on the naturalness of human languages gives very short shrift to the culturally determined side of languages. A proper theory of language and languages, though, must allow for both nature and culture and for their interaction in actual human languages, because no human system, including language, is or can be wholly natural. This observation extends also to the categories and kinds of human languages and cultures: they may be rooted in human nature but, because human nature encompasses the unnatural, so too should we expect the categories and kinds of human languages to be unnatural kinds rooted in nature. As Hurford (2011) argues at length, while the ability to acquire a human language can fruitfully be studied as a product of biological evolution, the languages themselves are also the products of cultural evolution. The entire phenomenon must therefore be viewed as the result of co-evolution of biology and culture.

### 2.3.1 *Inflectional classes and other purely morphological kinds*

I have tried to show that genders are best understood as unnatural kinds rooted in human nature. To repeat Corbett's observation: all genders have a semantic origin but there are very few 'strict semantic systems'. Nor are there purely morphological or phonological gender systems. Most gender systems are mixed and a little bit crazy in the same way that phonological rules can be crazy. But morphological categories freed from the syntactic tether can be a lot crazier than gender. Let's start with inflectional classes (Carstairs 1987), particularly the noun declensions and verb conjugations of traditional Latin grammar. Although Latin has three genders, it has at least five declensional classes of nouns (some have subclasses), which map onto genders in complex ways (Aronoff 1994). Even worse, there are four verb conjugation classes, which have no semantic or syntactic correlates of any type. Cross-cutting these is the active/passive distinction, which would correlate with syntax if it were not for the unfortunate existence of deponent verbs (Xu, Aronoff, and Anshen 2007), which by definition have active syntax but 'passive' morphology. Nor is Latin alone. All Romance languages retain conjugation classes for verbs, although all are somewhat simplified. And inflectional classes are not confined to Indo-European. Most Semitic languages have quite complexly distinct

inflectional classes of verbs. Dixon (1980) discusses a number of Australian languages whose morphology calls for a description in terms of inflectional classes. Historically, Indo-European inflectional classes are related to the existence of different theme vowels between the verb root and the desinences (Chantraine 1961, Ernout 1953). Semitic verb classes start out as syntactico-semanticly distinct (passive, causative, etc.) but don't remain that way (Aronoff 1994). In the Modern Aramaic dialect described by Hoberman (1992), the verb conjugation classes that are the reflexes of the ancient Semitic classes are distributed according to the number of syllables in the verb. In short, languages get crazy inflectional classes from what might have originally been natural phenomena, but the inflectional classes themselves are far from natural.

### 2.3.2 Morphemes

Morphemes are morphological patterns that have become completely unhinged. The term *morpheme* was originally coined (Aronoff 1994) to name a level of representation between morphosyntactic feature arrays and their phonological realization, in particular to single out cases where more than one morphosyntactic feature array maps onto the same set of phonological realizations. The example in the original discussion was the English perfect/passive participle, in which either the perfect or the passive maps onto a morphological function labeled  $F_{en}$ , which in turn is realized as one of many forms depending on the verb whose perfect/passive participle is being realized. Since both the perfect and the passive participle are realized by the same function, we need a name for the function and a general name for functions of this sort, which map one of many morphosyntactic feature arrays onto one of many realizations. In fact, I noted in the original discussion that the term could apply to any function that realizes morphosyntactic feature arrays but that 'it is morphemes like  $f_{en}$  that truly earn their name' (Aronoff 1994: 25).

Morphemes are functions within an incremental-realizational theory of morphology that map morphosyntactic representations onto phonological realizations. Their job is to link levels of representation and so they cannot be natural, by definition. They are arbitrary like Saussurean signs, just more complex. It is important to remember, though, that the morpheme is the whole function, not just the internal link or hidden level.

Using this definition of morphemes, there are four distinct types of morphemes, depending on the number of distinct feature arrays and the number of distinct realizations, as is schematized in Figure 2.1. First comes the question of **monovalent** vs **polyvalent** morphosyntax: does the morpheme in question realize one morphosyntactic feature array or a disjunction? Second is the question of **monomorphous** vs **polymorphous** realization: does the morpheme in question have one realization or is its realization context-sensitive? Here are examples of each type:

- **Monovalent monomorphous:** Kannada accusative suffix *-annu*. This suffix is invariant and unambiguous. All accusatives (including pronoun forms) and only accusatives are marked by *-annu* (Sridhar 1990).
- **Polyvalent monomorphous:** English *-Z*, which encodes plural in nouns, singular present in verbs, possessive, and a variety of contractions.
- **Monovalent polymorphous:** English past tense, which has a variety of realizations in ‘strong’ verbs, depending on the individual verb, and a default *-d*. Latin present and perfect stems are well-studied examples of this type.
- **Polyvalent polymorphous:** English *F<sub>en</sub>* is the classic case. It has two morphosyntactic values and maps onto many realizations, again depending on the individual verb. The Latin third stem, discussed at length in Aronoff (1994), is another.

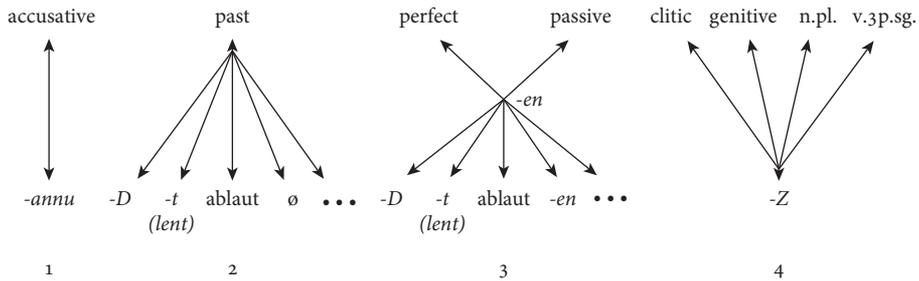


FIGURE 2.1 Graphic representations of the four examples

Since it was first introduced, however, the term *morphome* has come to be used especially for polyvalent polymorphous mappings; and it has been extended further to include all patterns where the morphological system has been set free from both syntactic and phonological tethers in the way that these mappings are. The paradigm cases come from the work of Martin Maiden (1992, 2005) on Romance verb paradigms, some of which prefigured the notion of the morphome. Maiden shows that, in a variety of Romance languages, certain cells of the verb paradigm may pattern together, even though there is no morphosyntactic or phonological motivation for the pattern. The most prominent of these is what he calls the ‘L-pattern’, which comprises all cells of the present subjunctive plus the first person singular present indicative cell. The ‘N-pattern’ consists of the all the singular cells and the third person plural cell of the present indicative and subjunctive. Neither of these can be justified on morphosyntactic grounds. They originated in Latin phonology but Maiden shows that they have persisted without phonological motivation and, in many cases, have spread far beyond their original distribution. These morphomic patterns determine the distribution of stems and even patterns of defectiveness (Maiden and O’Neill 2010). Maiden’s patterns are highly abstract entities, sets of cells in a paradigm. They are inherently irregular: a verb that shows L-pattern or

N-pattern allomorphy must be marked as having an unpredictable stem (or no stem) in precisely the cells that form the pattern. What makes a pattern of cells a morphomic category of the language is simply the fact that the pattern recurs, not any aspect of syntax or semantics or phonology. A category like this could never be natural in any substantive sense, which is why such categories are so important. The only way out of the conclusion that languages acquire at least some unnatural categories is to deny the validity of the phenomenon, to claim that these patterns are just figments of the analyst's imagination. Which is why Maiden's demonstration that the patterns persist, change, and grow over time is so important. The only sense in which these patterns can be said to be natural is if we believe that humans naturally construct patterns of this sort. This idea is not so far-fetched as it might at first appear. Human cultures share few common traits but one of the most striking is the human propensity for rule-governed behaviour, for creating structure. What is universal is not the structures, the categories, but the very act of categorization.

Now we are in a position to understand inflectional classes, which I defined in Aronoff (1994) as sets of lexemes whose members each choose the same set of inflectional realizations (Aronoff 1994: 64). Just like Maiden's cell patterns, the origins of inflectional classes are clear and motivated; and just like cell patterns, inflectional classes have escaped their semantically or phonologically motivated tethers. Again, the only way out, the only way to escape the conclusion that languages can have unnatural categories, is to deny the existence of inflectional classes. But such a denial is difficult, since inflectional classes have stared grammarians in the face for millennia, with no hope of reduction to some more motivated or natural kinds.<sup>4</sup>

## 2.4 Some sign language categories

A theory or hypothesis is valuable only if it helps us to understand something. What, then, is the value of claiming that it is natural for humans to construct unnatural categories, but that these categories are rooted in our minds/bodies and how we interact with the world, not determined in advance? I will now provide two sorts of phenomena from sign languages, each showing how the categories of these languages arise from systematization of the interaction of humans with the world.

### 2.4.1 Sign language verb agreement

The first phenomenon comes from the representation of sign language verb agreement systems. See Meir (2010) for a full discussion of what I can only summarize

<sup>4</sup> One reader suggests that inflectional classes can always be reduced to leading forms or theme vowels. Space prevents me from refuting this position here but much of Aronoff (1994) is an extended empirical demonstration that this suggestion will not work. Nor would this reduction vitiate what is most important about inflectional classes, which is that they are entirely language-particular.

here. Padden (1988) demonstrated that American Sign Language has verb agreement, but that only a subset of verbs, which she called agreement verbs, agree (with their subjects and objects). Later research found a number of other sign languages with similar systems of what one might call partial verb agreement and no sign languages with full agreement of the sort found in many spoken languages, where all verbs agree. The agreement in sign languages consists of first fixing referential loci, literally places in the signing space in front of the speaker where the referent of a noun phrase is conventionally placed, and then moving the verb sign from one locus to another. The referential loci can be used for other grammatical purposes as well and some sign languages have referential loci but do not show agreement. For those languages that do, the movement of the verb sign from one locus to the other constitutes morphological agreement and the verb sign inflects in the etymologically original sense of the Latin grammatical term *inflectere* ‘to bend’. The sign bends its shape, in this case its beginning and end point, to express agreement. For readers unfamiliar with sign languages, it is important to realize that most signs have movement, so that there is nothing odd about agreeing verb signs moving from one place to another; what is special about agreeing verbs is that where they move to and from is not fixed but varies depending on the referential loci of their arguments.

Meir (2002) showed that both the classification of verbs in sign language agreements systems and the direction and palm-facing of the movement are determined semantically: agreement verbs are verbs of transfer, the movement is from source to goal of the transfer, and the palm faces the syntactic object of the verb. Furthermore, all sign language agreement systems operate in the same way.

It should be obvious that this type of system is grounded in an iconic relation between the form of the verb and its meaning and that such an iconic relation could only arise in a sign language, as we showed in Aronoff, Meir, and Sandler (2005). There are no spoken languages in which only this subset of verbs shows morphological agreement, because spoken languages cannot represent this transfer from source to goal iconically. But it would be hard to claim that the category of agreement verbs is somehow given in advance. The category exists and is manifested morphologically and syntactically because of the iconic linguistic representational possibilities of the visual medium: it is ‘natural’ for users of a sign language to come up with a system in which transfer is represented iconically in this way and so to highlight the notion of transfer, but there is nothing innate about the category itself, which is why we never find evidence for it in spoken languages.

Further support for the contingent nature of verb agreement in sign languages is the fact that it does not appear in all sign languages. Comparing the two sign languages used in Israel that our research group has studied, Israeli Sign Language (ISL) and A-Sayyid Bedouin Sign Language (ABSL), Meir shows that ISL has developed a typical system of agreement for transfer verbs, while ABSL has not. In fact, ABSL shows no verb agreement at all. Thus, there is nothing necessary about the

category of sign language agreement verbs: the visual medium allows for such a category, maybe even promotes it, but it does not provide the category ready-made, and no spoken language uses such a category, because it would never come to the fore in a system where it is not representable.

The prototypical agreement verb in any sign language that is known to have the category is a ditransitive verb of transfer, the absolute archetype being GIVE. But once the category gains a foothold, it quite quickly accrues new members that do not fall within the semantic centre. In ISL, for example, the grammatical category of agreement verbs includes BE UNFAITHFUL TO and RESCUE, among many others. Table 2.1 contains a list of all ISL agreement verbs, from Meir (2002).<sup>5</sup> Because of the young age of the language, it is also possible to trace the history of some of these verbs. For older ISL and ASL signers, for example, the verbs FAX and TELEPHONE were not originally agreement verbs and some still do not use them as agreement verbs grammatically. For younger signers of both languages, however, both verbs show agreement. Each language will have its own list of agreement verbs, depending on a variety of factors. This sort of historical development is reminiscent of what happens to genders. They may start off as semantically motivated, but most lose at least part of this motivation. Importantly, the loss of motivation does not result in any weakening of the category. Quite to the contrary, genders remain entrenched, even when, as happens in a language like French, they become quite inscrutable semantically. Overall, linguistic categories are very conservative once they gain a foothold. This cumulative conservatism is a common trait of cultural and biological evolution: traits are slow to develop, but once they do, they do not easily disappear.

#### 2.4.2 *Object vs handling*

The second example from sign languages involves nouns that name concrete hand-held artefactual objects like combs, screwdrivers, and hats (Padden et al. 2013). In all cases, the objects in question are small enough to be easily held in some way by human hands. Some have handles or the dimensions of the object are such that it can be easily picked up and held in the hand.

For the above types of objects, the representational modality affords two iconic options: one is for the hands to represent how the artefact is handled; the other is for the hands to represent salient visual properties of the object itself. ASL has signs demonstrating both options. For example, a toothbrush is typically grasped between the thumb and the closed fingers, and thus could be represented by a closed handshape; alternatively, the sign could represent the shape of the brush itself, a long thin object, with an extended index finger. This latter form is the one actually

<sup>5</sup> The backwards verbs are those whose movement is from object to subject rather than the usual subject to object. See the references for detail.

**TABLE 2.1. A list of agreement verbs in ISL**

1st. Regular agreement verbs	2nd. Backwards verbs
<b>A.1. Di-transitive verbs</b>	<b>B.1. Di-transitive backwards verbs</b>
GIVE	GRAB
PAY~	RECEIVE/GET
SEND	MOOCH
SHOW	TAKE
THROW-TO	
<b>A.2. Monotransitive verbs</b>	<b>B.2. Monotransitive backwards verbs</b>
ANSWER*	ADOPT
APPROACH	CHOOSE
ASK*	COPY~
BE UNFAITHFUL TO	EXTRACT~
CALL (TELEPHONE)	INVITE
CATCH	RESCUE
DEFEND	SUMMON
EXPLAIN*	TAKE-ADVANTAGE-OF
FAX	
FEED~	
FINGERSPELL	
FIRE (a person)	
FORCE	
GUARD	
HATE (REBUFF)	
HELP	
INFLUENCE*	
INFORM*	
IRRITATE	
KEEP-AN-EYE-ON	
LIE-TO	
LOOK	
REVENGE	
SEE	
SHOOT	
SHOW-AFFECTION-TO	
SHOW-BELIEF-TOWARDS	
SHOW-CARE-FOR	
TEACH~	
TEASE	
TELL*	
TELL-STORY	
VIDEO-TAPE	
VISIT	
WARN*	
YELL-AT	

\* Verbs which may have an additional sentential complement.

~ Verbs which may occur both as monotransitive and as di-transitive verbs.

used in ASL. Other signs showing the object's dimension or extension include KNIFE, which shows the index and middle finger extended together in a cutting or sawing movement; FORK, which uses fingers for the tongs with a downward movement; and COMB, where the sign shows the tines as they move above the head in a circular movement. By contrast, other ASL signs refer to how the object is *handled*, as in HAMMER, to grip a hammer while hammering, or LIPSTICK, to hold the tube of lipstick while applying it on the lips. The contrast in form is whether the object is being held or whether the hands show dimensional aspects of the object as the task is performed. Both types have movement and location representing a typical task involving that object.

The ABSL lexicon includes vocabulary referring to categories, objects, locations, emotions, and abstract concepts. As we were collecting data on common objects in the Al-Sayyid Bedouin environment, we were surprised to discover a consistent lexical pattern among second- and third-generation signers with respect to artefacts that can be held by human hands. ABSL signers strongly prefer to use forms where a dimension of the object is represented by the fingers and the hand, the *object* type. When we compared how Israeli Sign Language produced forms for the same common objects, we found a different preferential pattern: ISL signers prefer forms where the signs show how the object is held by the hands, the *handling* type.

Because this preferential pattern had not previously been noted in the sign language literature, we developed a data set of pictures to show to signers of different sign languages, to see whether they would produce either of the preferential patterns we observed with ABSL and ISL signers.

The final set of 27 pictures include: six clothing items (shoes, trousers, hat, socks, gloves, and jacket), seven utensils (scissors, fork, knife, spoon, nail file, comb, and hairbrush), eleven hand tools (broom, paintbrush, toothbrush, rake, handsaw, screwdriver, hairdryer, mobile phone, vacuum cleaner, hammer, and mop) and three kinds of cosmetics (mascara, nail varnish, and lipstick). The pictures were assembled into a slide show, with a picture of each object on a single slide. In order to ensure elicitation of nouns rather than verbs, the slides varied with respect to the number of objects shown, e.g. one screwdriver or three combs. Signers were asked to identify the number and object, as a way to encourage naming rather than description of an event involving the object.

We found that signers of each of six different sign languages consistently preferred one or the other pattern for objects held by hand. The sign languages selected for this comparison of lexical type have no known historical relationship, and they vary in age. ABSL and ISL are the youngest sign languages of the group, both about 75 years old, having developed independently of each other in Israel. The remaining sign languages have existed since at least the nineteenth century: ASL, Danish Sign Language, Japanese Sign Language, and New Zealand Sign Language.

The sign languages vary in which lexical type is favoured; three (ABSL, ASL, and DSL) favour the object type and three (ISL, JSL, and NZSL), the handling type. They also vary with respect to the degree preferred. The pattern is not an exclusive one with one type used for all forms, but it is a *preferential* pattern.

Both the object and the handling pattern are iconic. One is iconic of the shape of the artefact and one is iconic of how the artefact is used. Both are salient and, in a sense, they compete. Even assuming that a language will consistently choose one pattern or the other for any given artefact, why do we find consistency across artefacts within a single language?

In the phenomenon we have discussed here, two strategies are available for the visual iconic representation of certain types of artefacts. Here is where the organized nature of language comes into play. At the level of the organization of individual lexical items, we can assume that the speakers of a language will agree, for any single artefact, that it should be represented in a constant manner, following either the object or the handling pattern. This is simple lexicalization. And this is what we find by and large (though occasionally a signer will use the handling pattern instead of the object pattern, when they wish to emphasize how the artefact is used). A priori, though, even assuming the lexicalization of individual noun signs, there is no reason to expect what we have actually found: that the signers of a language will agree on either the object or the handling pattern across the board for the majority of nouns for which there might be a choice. Our findings thus reveal that languages are not just opportunistic but organized. They take advantage of the resources that they are given and they use these resources to organize a system, even when, as in this case, there is no apparent reason to be quite so systematic. Importantly, though, there is no reason to believe that the organizational principles are given in advance, that they are natural in the sense under discussion in this article. Instead, the organization makes use of what is available to it, which in this case is the ways in which artefacts can be represented iconically in the visual mode.

## 2.5 Conclusion

One of the great victories of early modern linguistics was the acceptance of human languages as natural phenomena. I have tried to show here that, although language and languages are certainly natural to humans, we can be and have been led astray by analogizing too directly from nature to language. In particular, I have questioned the idea that the categories of language are given in advance in any strong sense. Instead, I have demonstrated that at least some linguistic categories can be profitably viewed as the result of the interaction between human minds/bodies and their environment. This is particularly dramatic in the case of the categories of sign languages, which can be very different from those of spoken languages, because they result from interaction in the visual realm. Other categories, genders for example, may begin in

real-world interaction and cognition, but can then take on a life of their own. Finally, there are morphological categories, morphemes, that have no reality outside the narrow system of the language itself. They lie at the core of language, the heart of darkness.

### **Acknowledgements**

Thanks to the organizers of the workshop that instigated the writing of this article, especially to Ana Luís for her great patience. Thanks to Ricardo Bermúdez-Otero, Greville Corbett, and Donca Steriade for their astute comments on an earlier version, and to the participants at the workshop. The research on sign languages that is reported here was supported by the National Institute on Deafness and Communicative Disorders of the United States National Institutes of Health through a grant to the University of California, San Diego.