Oxford Handbooks Online

Word Categories

Mark Smith

The Oxford Handbook of the Word
Edited by John R Taylor

Print Publication Date: Jun 2015 Subject: Linguistics, Lexicography
Online Publication Date: Mar 2014 DOI: 10.1093/oxfordhb/9780199641604.013.019

Abstract and Keywords

Linguists commonly reduce the grammatical behaviour of words to sets of properties associated with lexical categories such as noun and verb (‘parts of speech’). This chapter reviews a number of different approaches to the categorization of words and the problems associated with them. It is shown that problems arise when a word lacks properties associated with the category to which it has been assigned (‘subsective gradience’) and also when a word possesses properties associated with a category to which it is not assigned (‘intersective gradience’). It is also argued that the process of determining the set of categories within a language involves arbitrary and inconsistent choices that cannot be resolved on empirical grounds. Finally, it is argued that crosslinguistic variation in lexical categories and the properties associated with them undermine attempts to produce a crosslinguistically valid set of lexical categories.

Keywords: lexical categories, parts of speech, grammatical behaviour, subsective gradience, intersective gradience, crosslinguistic variation

10.1 Introduction

WORDS are often grouped into lexical categories (‘parts of speech’) on the basis of the types of entities that they denote. Thus, words may be grouped into a category of nouns if they denote things (e.g. *tomato, basket, piano*), a category of verbs if they denote events (e.g. *sing, run, eat*), a category of adjectives if they denote properties of things denoted by nouns (e.g. *red, hairy, loud*), and so on. In the Western linguistic tradition, this view of lexical categories has its roots in the Latin grammars of Priscian and Donatus and, ultimately, in the philosophy of Plato and Aristotle (Harris and Taylor 1997; Seuren 1998). However, modern linguists have repeatedly challenged this view. In particular, they have

observed that any given lexical category is liable to contain members that fail to denote the type of entity associated with the category (Baker 2003; Carstairs-McCarthy 1999). While many nouns do denote things (e.g. *basket*), some denote events (e.g. *explosion*), and some denote properties (e.g. *goodness*). In other cases, the matter is unclear. In the following examples, the adjectives *impossible* and *normal* have been used as nouns.

(1) a. There's no point trying to achieve the impossible.
b. Is economic stagnation the new normal?

It is by no means obvious whether these words denote things or properties, some combination of the two, or something else entirely.

The difficulty is compounded by the fact that some words don't seem to denote anything at all. In example (2), *that* is a type of word commonly termed a complementizer.

(2) John remembered that the plants needed watering.

(p. 176) As with other complementizers, such as *if* and *whether, that* has the grammatical function of marking out a subordinate clause (*the plants needed watering*) and signalling that it forms a part of the main clause (*John remembered*). The word *the* in this sentence is a type of word known as a determiner. Like *that*, it can hardly be said to denote anything at all. Instead, like other determiners (*a, this, some*, etc.), it indicates whether the object denoted by the following noun is definite or not. Thus, in (2), *the* suggests that John is thinking about a definite or specific set of plants in contrast to the following example which suggests that he is thinking about plants in general:

(3) John remembered that the plants needed watering.

As a final example, consider the following sentences:

(4) a. Mike's a hell of a cook!
b. What the hell is Dave doing that for?

Here, the noun *hell* does not denote hell. Rather, it combines with a determiner to form a noun phrase that conveys the attitude of the speaker. Thus, *a hell (of)* in (4a) suggests the speaker's admiration for Mike's ability as a cook, while *the hell* in (4b) indicates the speaker's contempt for Dave's behaviour. While some words, then, can be defined in terms of the types of entity that they denote, other words do not denote entities at all and must be defined in other terms, such as their grammatical function, abstract features such as definiteness, and the attitude of the speaker.

As well as grouping words into lexical categories, linguists often group the words within a given lexical category into a number of subcategories. Typically, the words within one subcategory will exhibit properties that are to some extent distinct from the properties of the words in other subcategories. Consider, for instance, interjections in English. According to Huddleston et al. (2002: 1361), the interjection category consists of 'words that do not combine with other words in integrated syntactic constructions'. This lack of

syntactic integration is most clearly demonstrated when interjections occur on their own as single word utterances (*Damn!, Ouch!, Oops!, Shh!*). Even when they occur as part of a longer utterance, they tend to precede a sentence rather than forming an integral part of it, or they are quoted in the form of distinct one-word utterances:

(5)
a. Damn! We've missed the bus again!
b. 'Oi!' yelled the police officer.
c. They all went 'Shh'.

In English, the interjection category contains a subcategory of ideophones. These are words such as *splash* and *bang* whose sound directly symbolizes their meaning (see Childs, this volume). The interjection category also contains a subcategory of expletive words such as *blast* and *damn*. These are typically derived from verbs and, as a result, (p. 177) they can occur with noun phrases in syntactically integrated utterances just as verbs can (Huddleston et al. 2002: 1361):

(6) Damn these mosquitoes!

Other interjections cannot be grouped into either of these subcategories (*ugh, oh, wow, oops, hey, ah, eh*, etc.). However, they resemble the interjections in the two subcategories insofar as they primarily function to express the mood of the speaker and often occur as syntactically unintegrated, one-word utterances.

Words can also be grouped into so-called mixed categories. These generally contain words that cannot easily be accommodated in standard lexical categories because they exhibit properties associated with two or more categories. In example (7):

(7) an amusing anecdote

the word *amusing* is like an adjective both because it denotes a property of an object denoted by a noun and because it immediately precedes a noun. At the same time, it is verb-like in that it is derived from the verb *amuse* and has the *-ing* suffix commonly found with verbs. As such, *amusing* is a typical member of the participle category whose members exhibit a mixture of adjectival and verbal properties.

Linguists also make use of ad hoc categories in order to capture the fact that particular words from different lexical categories have certain idiosyncratic properties in common. Consider the italicized words in the following example:

(8)
a. What *ever* did you do that for? (adverb)
b. What *the hell* did you do that for? (noun phrase)
c. What *on earth* did you do that for? (prepositional phrase)

In these sentences, the italicized words modify the interrogative pronoun *what* and express the angry or incredulous mood of the speaker. Because they have these properties in common, Huddleston (2002: 916) groups them together into a lexical

category he terms ‘emotive modifiers’. At the same time, he recognizes that they can also be described as belonging to a wide range of other categories, such as adverb, noun phrase, and prepositional phrase. Similarly, Pullum and Huddleston (2002a: 823–8) postulate an ad hoc category they term ‘negatively-oriented polarity-sensitive items’ (or NPIs for short). The members of this category all prefer to occur in sentences with a negative meaning. For instance, Sinclair (2004) analysed the thirty occurrences of the verb *budge* in a corpus of 20 million words and found that the word always occurs in sentences denoting a failure to move, as in the following:

(9)
a. But Mr Volcker has yet to budge …
b. … but he refuses to budge on design principles he knows to be …
c. I determined not to budge from it until closing time.

(p. 178) There were no examples such as the following in Sinclair’s corpus:

(10) I pushed it and it budged.

Similarly, Taylor (2012: 60) observes that the NPI (and quantifying phrase) *much of* can only occur in negative sentences when it functions to evaluate the object denoted by the following noun:

(11)
a. I’m not much of a cook.
b. *I’m much of a cook.
c. They don’t think I’m much of a cook.
d. *They think I’m much of a cook.

The NPI category is a typical ad hoc category in that it groups together words and phrases from a wide range of different lexical categories all of which exhibit an idiosyncratic property not associated with other lexical categories.

It is also the case that lexical categories can be grouped into larger supercategories. For instance, a distinction is often made between content word categories (e.g. nouns, verbs, adjectives, adverbs), whose members denote types of entities, and function word categories (e.g. determiners, complementizers, prepositions), whose members have a primarily grammatical rather than semantic function. However, there are a number of problematic examples which suggest that it is difficult to draw a clear distinction between the two supercategories. It has often been observed that some content word categories contain function words. Consider the following sentences from Leech and Li (1995: 190):

(12)
a. …the question would seem to be a legitimate *one* …
b. He looked a mournful *man*.
c. ‘I became a political *being* for the first time in my life,’ she said.

The italicized nouns provide little new semantic information. Instead, the semantic information is given primarily by the adjective modifying the noun. Rather than having a semantic function, the noun's role is mainly grammatical—it supplies the obligatory head noun required by noun phrases in English. In this way, 'dummy nouns' are examples of function words that belong in a content word category, namely the noun category. A further problem is that some function word categories contain content words. The preposition category is often regarded as a function word category (see e.g. Chomsky, 1995) but most of its members are clearly semantic insofar as they can denote temporal and spatial relations (e.g. *during, after, between*). Moreover, certain prepositions are content words in some contexts and function words in others (Hudson 2000). In (13a), *by* has a spatial meaning, comparable to that of phrases such as *next to* or *in proximity to*. In contrast, in (13b), *by* has the grammatical function of marking the attacker as the agent of the activity.

(13)
a. Dave stood by the fire.
b. Sue was being pursued by her attacker.

(p. 179) Such observations cast doubt over whether a clear distinction can be maintained between content word categories and function word categories.

In contemporary linguistics, lexical categories underpin a view of language that has been variously labelled the 'slot and filler model' (Sinclair 1991: 109) and the 'dictionary plus grammar model' (Taylor 2012: 44). According to this view, lexical categories enable words to be combined into phrases and sentences on the basis of a set of grammatical rules. A grammatical rule might stipulate that a determiner (e.g. *the*), an adjective (e.g. *sleepy*), and a noun (e.g. *cat*) can be combined together into a noun phrase (*the sleepy cat*). Another rule might stipulate that a noun phrase (e.g. *the sleepy cat*) and a verb phrase (e.g. *yawned*) can be combined together into a sentence (*The sleepy cat yawned*). Basing grammatical rules around lexical categories in this way allows the number of grammatical rules that make up a grammar to be greatly reduced. If English words can be grouped into eight or nine lexical categories, as proposed by e.g. Aarts (2007) and Huddleston and Pullum (1984), this implies that a grammar of English need only comprise eight or nine sets of rules, one for each category, in order to be able to describe the grammatical behaviour of all of the words in all of its lexical categories. If, however, English words cannot be grouped into lexical categories, then the grammar would need to comprise a distinct set of rules for each of the hundreds of thousands of words in the language in order to be able to describe their grammatical behaviour. In this way, lexical categories make possible a simple and elegant description of how words combine into phrases and sentences. Without lexical categories, grammar would become hopelessly complex and chaotic.

Yet while lexical categories are of fundamental importance in grammar, much recent work has argued that they are highly problematic. In particular, these studies have suggested that the grammatical behaviour of words is far too complex to be captured by the properties associated with lexical categories (Croft 2001; Crystal 1967/2004;

Culicover 1999; Francis 1993; Gross 1994; Sinclair 1991, 2004; Taylor 2012). They have also argued that determining the lexical categories of a language and the properties associated with them is inevitably a highly arbitrary process (Croft 2001; Plank 1984; Smith 2010, 2011). The remainder of this chapter gives an overview of this recent critical work. It concludes that there are a number of deep and intractable problems associated with lexical categories, and that lexical categories fail to provide an adequate description of the grammatical behaviour of words.

10.2 Subsective and intersective gradience in categories

It is rarely the case that all the words in a lexical category will exhibit all the properties associated with that category. Indeed, it is rarely the case that all the words within a category will share even a single property. Consider, for instance, adjectives in English. Properties associated with the category include the ability to occur in prenominal and (p. 180) predicative positions, the ability to take the prefix *un-*, and the ability to be intensified and graded (Aarts 2007).

(14)

a. a happy woman	(prenominal position)
b. She is happy	(predicative position)
c. very happy	(intensification)
d. happy/happier/happiest	(gradedness)
e. unhappy	(*un*-prefixation)

Many adjectives are unlike *happy* in that they do not exhibit all of these properties. Table 10.1 is adapted from Aarts (2007: table 8.2).

Table 10.1 Adjective criteria

	Prenominal	**Predicate**	**Intensification**	**Gradedness**	***un*-prefix**
happy	+	+	+	+	+
thin	+	+	+	+	–
afraid	−	+	+	+	+
alive	−	+	+	?	−
utter	+	−	−	−	−
galore	−	−	−	−	−

In Table 10.1, there is no single property which is exhibited by all six adjectives. Moreover, even a word such as *happy*, which displays all five of the listed properties, does not exhibit all the properties associated with English adjectives. For example, as Ferris (1993) notes, *happy* generally does not occur postnominally (e.g. **the winner happy*) in contrast to adjectives such as *galore* which do (e.g. *restaurants galore*; also *for reasons unknown, the body beautiful*). This tendency for the words within a lexical category to exhibit some but not all the properties associated with that category is known as subsective gradience (Aarts 2007). Subsective gradience can be linked to the notion of prototypicality (Rosch 1978). Specifically, some words will be better, more prototypical examples of their category than other words because they exhibit more of the properties associated with it. *Happy* is a prototypical adjective because it exhibits so many adjectival properties. In contrast, *utter* is not a prototypical adjective because it exhibits so few adjectival categories.

As a further example of subsective gradience, consider nouns in English. A key characteristic of nouns is their ability to occur either as the subject of a verb (e.g. *professor* in the following example) or as the object of a verb (e.g. *students*):

(15) The professor thanked the students.

However, a number of nouns only occur inside prepositional phrases (e.g. *sake* in *for the sake of, accordance* in *in accordance with, dint* in *by dint of*), and while most nouns (p. 181) can be marked as plurals (e.g. *students* in the above example), many nouns cannot (e.g. *hydrogen, machinery, furniture*). Further evidence of subsective gradience is provided by the category of English determiners. Culicover (1999) analysed the behaviour of eighteen determiners (e.g. *every, this, each, our*) in terms of six different properties. He observed that none of the determiners exhibits all six properties even though each property is exhibited by at least one of them. For instance, some determiners such as *this* can appear without an overt head whereas others such as *every* cannot (Culicover 1999: 62):

(16)
a. I'll take this.
b. *I'll take every.

Some determiners such as *all* can shift to the right whereas others such as *our* cannot:

(17)
a. All the women have gone.
b. The women have all gone.
c. Our women have gone.
d. *Women have our gone.

Some determiners such as *many* can occur between a determiner and a noun just like an adjective, whereas other determiners such as *both* cannot.

(18)
a. Water conditioning is just one of the many uses of salt.
b. *I thought the both entries were very impressive.

The English adjective, noun, and determiner categories are not unusual in exhibiting subsective gradience. Instead, most lexical categories in most languages exhibit such variation to some extent (Aarts 2007; Bhat 1994; Croft 2001; Crystal 1967/2004; Culicover 1999; Hunston and Francis 2000; Plank 1984; Smith 2011).

Moreover, the words within a given lexical category are liable to vary not only in terms of subsective gradience but also in terms of intersective gradience. Intersective gradience concerns the propensity of the words within a given lexical category to exhibit properties associated with other lexical categories. Consider, again, the adjectival properties listed in Table 10.1. Like adjectives, nouns can also occupy a prenominal position (e.g. *apple* in *apple pie*). Similarly, nouns often occur predicatively (e.g. *teacher* as in *John is a teacher*) and can be modified by *very*, as the following example from the British National Corpus demonstrates:

(19)
Indeed, the very notion of student access implies curricular foundations.

Moreover, adverbs can be modified by *very* (e.g. *quickly* as in *very quickly*) and can be graded as the following example from the British National Corpus demonstrates:

(20)
Familiarity increased most quickly, and is now highest, with well-educated people.

(p. 182) Finally, many verbs can take an *un*-prefix (e.g. *do* as in *to undo*). Moreover, adjectives often exhibit properties associated with other categories. Thus, determiner marking is typically regarded as a nominal property (although pronouns, proper names, and some common nouns such as *accordance* and *fatherhood* typically occur without a determiner). Nevertheless, adjectives often occur with determiners as the following examples from Ferris (1993) demonstrate:

(21)
a. The older a violin is, the more valuable it is supposed to be.
b. Martin is now the thinnest he's ever been.
c. Father came back $500 the poorer.

Moreover, some adjectives can take noun phrase complements (e.g. *worth* in *It was worth a hundred pounds* or *like* in *She is so like her brother*), even though this is a property which is typically associated with prepositions and verbs.

While the words of a given lexical category will often be able to display properties associated with other lexical categories, there will typically be a range of other properties

that they cannot display without being converted into a different lexical category. For the adjective *kind* to occur like a noun as the subject of a verb, the category converting suffix *-ness* must be added to it to indicate that it is now functioning as a noun:

(22) Your kindness touched me deeply.

Similarly, for the verb *develop* to occur like a noun as the object of a verb, the category converting suffix *-ment* must be added to indicate that it is now functioning as a noun:

(23) The entire organization requires substantial development.

Nevertheless, certain words seem not to respect the limits of intersectively gradient behaviour associated with their category. In particular, they seem to be able to display an extremely wide range of properties without having a category converting suffix added to them. Consider the word *spare*. It can occur as an adjective, a verb, or a noun without the addition of a category converting suffix:

(24)
a. They couldn't even spare me a dime. (verb)
b. Have you got a spare cigarette? (adjective)
c. There is a garage nearby that can sell you spares. (noun)

Note that the *-s* suffix in (24c) is not a category-converting suffix; its function is to mark *spare* as plural rather than to signal that it has been converted into a noun. Moreover, (p. 183) this lack of category converting suffixes can even render the lexical category of *spare* ambiguous, as the following examples demonstrate:

(25)
a. Have you got a spare?
b. Have you got a spare one?
c. Have you got any spares?

It can be argued that *spare* in (25a) is an adjective modifying an ellipsed noun, as suggested by the possibility of adding the dummy noun *one* (25b). However, it can also be argued that *spare* in (25a) is a head noun in a noun phrase, as suggested by the possibility of adding the plural marker *-s* (25c). One conclusion might be that *spare* is so intersectively gradient that it straddles different categories and blurs the boundaries between them.

Subsective and intersective gradience suggest that we cannot restrict membership of a lexical category to those words exhibiting all and only the properties associated with the lexical category. Restricting membership in this way would exclude a large proportion of the words typically included in a given category. As a result, it is necessary to take a more flexible approach to category membership. In much recent work, this has been achieved by means of the Best Fit Principle. This states that we should assign a word to whichever category it possesses the most properties of (Aarts 2007; Crystal 1967/2004; Plank 1984). If *utter* exhibits a single adjectival property and no properties associated with other categories, then it will be assigned to the adjective category. If a word exhibits some

properties associated with the verb category, then it still might be assigned to the noun category if it exhibits more noun properties than verb properties. Consider the following example (from Huddleston 1984: 307):

(26) These killings must stop.

The gerund *killings* exhibits two verb properties: it is derived from the verb *kill* and contains the suffix *-ing* typically associated with verbs. At the same time, it exhibits three noun properties; it occurs as the subject of a verb, contains the plural marking suffix *-s,* and is marked by the determiner *these*. Consequently, *killings* would be assigned to the noun category on the basis of the Best Fit Principle. In this way, then, the Best Fit Principle allows lexical categories to accommodate intersectively gradient words.

The Best Fit Principle is problematic in a number of respects (Smith 2011). Because the principle is so tolerant of subsective and intersective gradience, it can sanction lexical categories like the English adverb category which contains such a miscellaneous set of members that it risks becoming incoherent. While adverbs are often defined as words that modify verbs and adjectives, the English category contains words such as *too* that modify adjectives and adverbs but not verbs, and words such as *upstairs* that modify verbs and nouns but not adjectives or adverbs:

(27)
a. It was too *silly* to contemplate. (adjective)
b. Sara had eaten her food too *quickly*. (adverb)

(p. 184) **(28)**
a. They *ran* upstairs as fast as they could. (verb)
b. The *room* upstairs is no longer vacant. (noun)

Moreover, adverbs such as *almost* can modify an astonishing range of different kinds of words and phrases (Pullum and Huddleston 2002b: 562):

(29)
a. They almost *suffocated*. (verb)
b. The article was almost *incomprehensible*. (adjective)
c. She almost *always* gets it right. (adverb)
d. Almost *all* the candidates failed. (determiner)
e. They are almost *without equal*. (prepositional phrase)
f. She read almost *the whole book* in one day. (noun phrase)

Given this bewildering variety of grammatical behaviour, it is difficult to pinpoint exactly what kind of word an adverb is. Nevertheless, because the Best Fit Principle allows for such variation, it is able to generate this kind of ragbag category.

The Best Fit Principle gives rise to a number of other problems. For instance, the principle would assign a gerund that exhibited five noun properties and six verb properties to the verb category. In so doing, however, it would fail to reflect the fact that the gerund is almost as strongly associated with the noun category as it is with the verb category, and the fact that it may exhibit more nominal properties than many words

categorized as nouns. A model employing the Best Fit Principle would also assume that all the properties associated with a category are equally diagnostic of that category. In many instances, such an assumption is contentious. Aarts's model assumes that the ability to take a nominal suffix and the ability to occur as a subject are equally diagnostic of nounhood. However, nouns lacking nominal suffixes are often highly prototypical (e.g. *shoe, boat*) while nouns that cannot occur as subjects are not at all prototypical (e.g. *dint, sake, accordance*).

The Best Fit Principle also assumes that any given property will be associated with only a single category. But why, for instance, should we regard the ability to occur in predicate or prenominal position as exclusively adjectival properties and not also as nominal properties, given that nouns also occur in these positions?

(30)
a. John is a teacher
b. a piano factory

Similarly, why should we regard the ability to occur with a determiner as a nominal property when adjectives can also occur with determiners, as (21) above demonstrates? Moreover, it is often unclear whether a given word should be regarded as exhibiting a particular property or not. For instance, it is natural to assume that pronouns lack the ability to occur with determiners because pronouns and determiners so rarely co-occur. Nevertheless, an internet search will return perfectly legitimate instances of pronouns occurring with determiners, as in the following: (p. 185)

(31)
What's inside can make all the difference in the you you turn out to be.

Should we regard this example as evidence that pronouns can occur with determiners? If we do, we overlook the important fact that determiners regularly occur with common nouns but only rarely with pronouns. But to assume that pronouns do not occur with determiners leaves us unable to account for examples such as (31). While, then, the Best Fit Principle does enable lexical categories to accommodate subsectively and intersectively gradient words, it does so only at the cost of simplifying and distorting the complex grammatical behaviour of words.

10.3 Quirky properties

Many of the grammatical properties exhibited by words cannot be explained in terms of properties associated with lexical categories. These ‘quirky’ properties are exhibited by only a handful of words or even by a single word. For instance, *at* can occur with *night* and *night-time* but not with *day* and *daytime* (Taylor 2012: 97–8):

(32)
a. at night
b. at night-time
c. *at day
d. *at daytime

This contrast cannot be explained in terms of lexical category differences because *night, night-time, day*, and *daytime* are all nouns. It is simply a quirky property of *at* that it occurs with certain words but not others. Similarly, while *burning* can occur with *ambition*, it cannot occur with *row*. The reverse is true for *blazing*:

(33)
a. a blazing/*burning row
b. a burning/*blazing ambition

This contrast cannot be explained in terms of differences in lexical category because *blazing* and *burning* are both participles and *row* and *ambition* are both nouns. It is simply a quirky property of *row* and *ambition* that they combine with certain participles but not others.

There can also be quirky restrictions on the kinds of words that occur in a particular grammatical structure. Consider the following examples:

(34)
a. my love of strawberries
b. my hatred of strawberries
c. *my hate of strawberries

(p. 186) *Love* and *hatred* can occur in a grammatical structure comprising a noun modified by a possessive pronoun (*my*) and an *of* phrase (*of strawberries*). In contrast, *hate* cannot occur in this structure even though it is a noun like *love* and *hatred*. As a result, (34c) cannot be explained in terms of lexical category differences. It is simply a quirky property of *hate* that it is barred from a grammatical structure that other semantically similar nouns can occur in. A further example of quirky restrictions was observed by Francis (1993) in an analysis of the occurrences of the adjective *possible* in the Bank of English corpus. This demonstrated that *possible* can occur between a superlative adjective and a head noun, and following a superlative adjective and head noun:

(35)
a. I have the *strongest possible* belief in him …
b. … in the *best way possible* by beating them.

The only other adjective to occupy such positions in the corpus is *imaginable*. Francis also observes that *possible* can occur in the rightmost gap of *as … as …* structures.

(36)
a. She tried to visit as little as possible.
b. Try to stay as relaxed as possible.

A wide range of adjectives, adverbs, and quantifiers can occur in the leftmost gap of *as … as …* structures (e.g. *soon, quick*). However, it is difficult to substitute any other adjectives for *possible.* Examples such as the following are not attested in the corpus:

(37)
She tried to visit as little as??feasible/??imaginable/??plausible/??likely.

Other environments in which *possible* is commonly found, but where other adjectives hardly ever occur, include adjunct phrases such as *where/wherever possible, when/whenever possible*, and *if possible*. Clearly, the distinctive behaviour of *possible* cannot be explained in terms of its lexical category because other adjectives cannot occupy the same grammatical positions. It is simply a quirky property of *possible* that it can occur in a number of grammatical structures from which other adjectives are barred.

A word can also exhibit its own quirky restrictions on the properties associated with lexical categories. The word *go*, for instance, typically occurs as a verb but it can exhibit nominal properties when it occurs as an argument of *have* and other light verbs:

(38)
a. How many goes did you take then?
b. They were so beautiful that I decided to have a go at growing them.
c. She just knew she could make a go of it!
d. Phoebe decided to give it a go.

As these examples (from the British National Corpus) illustrate, *go* can exhibit various nominal properties such as plural marking and occurring in argument positions with quantifiers and determiners. Crucially, when *go* occurs as an argument of *have* and (p. 187) other light verbs, its nominal properties are highly restricted. For instance, it can occur with the indefinite article but not with the definite article:

(39) *I decided to have the go at growing them.

When it combines with light verbs such as *have*, it can occur in a direct object phrase but not as subject of a passive verb:

(40) *A go was had at the referee by the players.

Nominal *go* is quirky insofar as it typically occurs as the argument of a restricted set of light verbs. Even when it occurs in this construction, there are quirky restrictions on the nominal properties that it exhibits. These restrictions cannot be explained in terms of the fact that nominal *go* is a noun since other nouns typically do not exhibit such restrictions.

While we cannot explain quirky properties in terms of lexical categories, a number of studies have suggested that they may be subject to semantic restrictions. For instance, Stubbs (2002) analysed approximately 40,000 occurrences of *cause* as both a noun and verb in the Bank of English corpus. He found that all fifty of the words that co-occurred most frequently with *cause* had unpleasant connotations (*cancer, pain, trouble*). Stubbs also analysed 400 occurrences of *undergo* from the Cobuild corpus and found that most of the twenty words that co-occurred most frequently with it were either medical in nature (*heart, medical, surgery*) or denoted the serious or involuntary nature of events (*major, forced, required*). Yet while some quirky properties can be explained in semantic terms, it is clear that others cannot. The contrast between *my hatred of strawberries* and *my hate of strawberries* cannot be motivated semantically, because *hate* and *hatred* are synonymous. Gross (1994: 250–51) makes a similar point with regard to quirky properties in French. He observes the following contrast between *concerner* and *regarder*:

(41)
a. Cette affaire concerne Max.
b. Cette affaire regarde Max.
c. Max est concerné par cette affaire.
d. *Max est regardé par cette affaire.

It is difficult to motivate this contrast semantically because *concerner* and *regarder* have similar meanings. Both (41a) and (41b), for instance, can be translated as ‘This affair concerns Max’.

Like subsective and intersective gradience, quirky properties suggest that the behaviour of a word can only be partially captured by the set of properties associated with lexical categories. While subsective gradience shows that a word may not exhibit all of the properties associated with its category, and intersective gradience shows that a word may exhibit some of the properties associated with a different category from its own, quirky

properties demonstrate that a word may exhibit properties (p. 188) not associated with any category at all. In each case, the net effect is the same—the three phenomena all imply that lexical categories and their associated properties are only partially successful in capturing the behaviour of words. While some of the behaviour of words such as *possible* and *go* can be captured in terms of the properties associated with their lexical categories, large parts of their behaviour cannot be. As a result, any comprehensive description of their behaviour must specify not only the set of properties associated with their lexical category but also how they are subsectively and intersectively gradient and what quirky properties they exhibit. Words such as *possible* and *go* are not unusual in this respect. In fact, when we study actual language usage, it can be difficult to find a word which does not exhibit some subsective or intersective gradience or some quirky properties (Gross 1994; Partington 1998; Sinclair 1991, 2004; Stubbs 2002; Taylor 2012). Francis summarizes the situation as follows:

> If we take any one of a huge range of the more frequent words in English, and examine its citations en masse, it will emerge that it, too, has a unique grammatical profile, which certainly cannot be encapsulated by calling the word in question an adjective or a noun or a preposition.
>
> (Francis 1993: 147–8)

10.4 Lumping and splitting categories

Words within a lexical category do not vary at random. Instead, within any category, groups of words can often be discerned which cluster together insofar as they exhibit similar patterns of properties. As noted in section 10.1, these groups are often referred to as subcategories. Nouns in English, for instance, are often divided into four subcategories: common nouns, pronouns, proper names, and gerunds, each of which is associated with a distinct set of properties. For instance, common nouns often occur with determiners and adjectives (*a happy child*) but are not marked for case, while pronouns rarely occur with determiners and adjectives but can exhibit case marking (*he* vs. *him*). Moreover, groups of words exhibiting distinct sets of properties can also be discerned within each of the subcategories. The pronoun category can be split into possessive (e.g. *your*), demonstrative (e.g. *these*), personal (e.g. *him*), and reflexive (e.g. *himself*) categories. While possessive and demonstrative pronouns can occur before a noun, personal and reflexive pronouns tend not to. These categories can then be split even further. The personal pronoun category can be split into personal pronouns that occur only as subjects (e.g. *I*), those that occur only as objects (e.g. *him*), and those that occur both as subjects and objects (e.g. *you*).

In this way, then, categories can be split into ever smaller categories each of which exhibits its own distinctive set of properties. If a very wide range of properties are considered, then splitting can continue until a large open-class category has been split into (p. 189) hundreds of distinct categories each containing only one or two words. Gross (1994) analysed 12,000 French verbs in terms of a set of 300 properties. He observed that only about seventy French verbs enter into the construction N V à N where the second noun denotes a human, as in the following:

(42)
a. Max obéit à Bob.
b. Max pense à Bob.

Of these seventy verbs, only around twenty allow the second noun to be converted into a pronoun occurring before the verb:

(43)
a. Max lui obéit.
b. *Max lui pense.

Using a set of 300 properties such as these, Gross (1994) split the 12,000 into 9,000 distinct categories each exhibiting a unique combination of properties. Very few of these categories contained more than two verbs. Gross predicted that if the number of properties had been increased still further, no two verbs would have exhibited exactly the same set of properties; consequently, no two verbs would have belonged to the same category. Each category would have had only one member.

Splitting increases the number of categories and decreases the number of words in each category. As splitting renders categories progressively smaller, the members of those categories become increasingly homogeneous in terms of the properties that they exhibit. As a result, the set of properties associated with each category would provide an increasingly precise description of the grammatical behaviour of its members (Croft 2001; Plank 1984). Yet while splitting would result in a precise description of word behaviour, it would also result in a complex and uneconomical model of the grammar, one which would lack information about any categories larger than the smallest subcategories. It would fail to capture the fact that words within larger categories, such as nouns, do tend to share some common properties.

A model that featured a small number of large categories (categories such as nouns, verbs, and adjectives) would clearly be simpler and more economical. However, since each category would contain an extremely heterogeneous collection of words (e.g. the noun category would lump together common nouns, pronouns, gerunds, and proper names), the set of properties associated with that category would provide only a very imprecise description of the grammatical behaviour of its members (Crystal 1967/2004; Plank 1984). Moreover, the set of properties would fail to capture the similarities in grammatical behaviour between members of smaller subcategories. Thus, the single set of properties associated with the noun category would fail to capture the fact that words within smaller categories such as gerunds tend to share a distinct set of properties.

There is also the danger that lumping, like splitting, can be carried to extremes. As noted in section 10.2, if a model of lexical categories employs a flexible approach to category membership, as sanctioned by the Best Fit Principle, it can allow categories to (p. 190) include subsectively and intersectively gradient words. Thus, Aarts's (2007) model can lump common nouns, pronouns, proper names, and gerunds into a single noun category, even though these words exhibit contrasting sets of properties. However, because Aarts's model is so tolerant of subsective and intersective gradience, it can also lump nouns, verbs, and adjectives into a single supercategory, by simply attaching all the properties associated with nouns, verbs, and adjectives to the supercategory. Because any noun, verb, or adjective would be likely to display more of the properties associated with this supercategory than any other category, they would be assigned to this supercategory on the basis of the Best Fit Principle. Aarts (2007: 436–7) attempts to guard against this danger by stipulating that categories should not be 'too heterogeneous'. However, any attempt to specify exactly how much heterogeneity is acceptable is bound to be arbitrary and subjective (Smith 2011). For instance, we may feel that a supercategory that lumped together nouns and adjectives would be far too heterogeneous. Yet Donatus's grammar of Latin did precisely this and was the most widely accepted and influential grammar until the 11th century (Seuren 1998). At least according to medieval tastes, such a supercategory would not be too heterogeneous.

The grammatical behaviour of the words in a language can be described using either a small set of large categories or a large set of small categories. The former yields an economical but imprecise model of the grammatical behaviour of words, but fails to capture generalizations regarding smaller categories. In contrast, the latter approach yields an uneconomical but precise model of the grammatical behaviour of words, but fails to capture generalizations regarding larger categories. Because both approaches are successful in capturing some aspects of the data but flawed insofar as they fail to capture other aspects, it is difficult to see how either approach can be regarded as 'correct' or even better than the other. Each approach has its different strengths and weaknesses.

10.5 Arbitrary and inconsistent categories

We have seen that no approach to categorizing the words in a language can be regarded as inherently correct. This raises the concern that any choice between different approaches will be little more than an arbitrary matter of taste. Croft (2001) makes the point that in many perennial linguistic debates, such as whether Japanese has one or two adjective categories, or whether North American languages such as Nootka distinguish nouns and verbs, each side inevitably presents a model which succeeds in capturing some aspects of the data but fails to capture other aspects. In order to present their model as empirically correct, each side emphasizes the importance of the data that their model gets right while dismissing the relevance of the data that it fails to capture. As an example of this problem consider the word *piano* in the phrase *a piano factory*. There has

been a long-running debate over whether such a word should be categorized as an adjective or a noun (Smith 2010). (p. 191) That *piano* occurs before a noun and modifies its meaning both suggest that it is an adjective, as does the fact that it cannot support plural marking (e.g. **a pianos factory*). However, the fact that it denotes an object and the fact that it has the same form as a word that typically occurs as a noun (e.g. *piano*) both suggest that it is a noun. Consequently, in order to argue that it is an adjective, it is necessary to emphasize the importance of its adjectival properties while dismissing the relevance of its nominal properties. Conversely, in order to argue that it is a noun, it is necessary to emphasize the importance of its nominal properties while dismissing the relevance of its adjectival properties. Both sides misrepresent their own model as the only empirically correct one when in fact both are empirically correct in certain respects and empirically flawed in other respects. While the choice between the two is presented by each side as a choice between an empirically correct model and an empirically incorrect model, the choice cannot be decided on empirical grounds but is ultimately, as Croft notes, little more than a matter of personal taste.

Croft (2001) also raises the concern that many of the decisions that inform models of lexical categories are not only arbitrary but also inconsistent. Suppose, for instance, that we have arbitrarily chosen to regard *piano* in *a piano factory* as a noun rather than an adjective on the basis that it denotes a kind of object. If we choose to privilege such a semantic criterion over other criteria in this way, then we should at least apply this criterion consistently when we come to make other decisions about other lexical categories. We should decide that *examination* in *the examination of the patient* is a verb rather than a noun on the basis that it denotes an event rather than a thing. In fact, linguists are rarely consistent in this way because such consistency would typically lead to odd and counterintuitive categorizations (Croft 2001, 2007). Instead, linguists typically apply certain sets of criteria in deciding the category of certain words and different sets of criteria in deciding the category of other words. Thus, one might argue that *piano* in *a piano factory* is a noun because of its semantic properties while simultaneously arguing that *examination* in *the examination of the patient* is a noun because of its grammatical properties. Models of lexical categories, then, are typically a product of a series of decisions that are both arbitrary and inconsistent.

Croft's (2001) own model of nouns, verbs, and adjectives avoids such inconsistency because it consistently defines these categories in terms of a combination of semantic and pragmatic properties. Specifically, the model stipulates that a prototypical noun exhibits the semantic property of denoting an object and the pragmatic function of reference, a prototypical verb exhibits the semantic property of denoting an action and the pragmatic function of predication (i.e. saying something about the object denoted by a noun phrase), and a prototypical adjective exhibits the semantic property of denoting a property and the pragmatic function of modification (i.e. specifying the meaning of the object denoted by a noun). Words which exhibit a different combination of semantic properties and pragmatic functions are left uncategorized by Croft's approach. Consider again *piano* in *a piano factory*. This word combines the semantic property of denoting an object with the pragmatic function of modification. As such, it does not qualify for

inclusion in any of Croft's categories. What is more, Croft is not able to state what category it does belong to (Croft 2010; Smith 2010). We could, for instance, categorize the word as a non-prototypical noun on the basis that it exhibits the semantic property of (p. 192) denoting an object—a property associated with nouns—but the pragmatic function of modification—a property associated with adjectives. However, doing this would involve the arbitrary decision that it is the semantic property of a word rather than its pragmatic function which determines its category. Such a decision might commit us to the view that *examination* in *the examination of the patient* is a verb rather than a noun on the basis that it exhibits the semantic property of denoting an action associated with verbs. Alternatively, we could decide that *examination* is a noun on the basis that it exhibits the pragmatic function of reference that is associated with nouns. However, this would be inconsistent with our prior decision to define nouns in terms of semantic properties and consequently to categorize *piano* in *a piano factory* as a noun. In this way, then, attempting to categorize words that do not belong to the prototypical noun, verb, and adjective categories necessitates arbitrary and inconsistent decisions. These, however, are precisely the kinds of decision that Croft is determined to avoid. Unfortunately, by avoiding these decisions, Croft's approach is incapable of categorizing all those words (prenominal nouns, gerunds, action nominals, infinitives, participles, etc.) that do not belong in his prototypical categories (Smith 2010). In effect, Croft deals with subsectively and intersectively gradient words by simply excluding them from his model of lexical categories.

10.6 Categories from a cross-linguistic perspective

Just as the properties associated with particular lexical categories vary within languages, so too they vary between languages. A given property might be subsectively variant across languages insofar as the property is exhibited by words in a particular category in some languages but not by the words in a particular category in other languages. In many languages, for instance, nouns exhibit number, gender, and case marking. Vietnamese, however, lacks these features (Croft 2001), so these properties are irrelevant for Vietnamese. Similarly, nouns in many languages take determiners, but this property is irrelevant to Latin, since Latin lacks determiner marking. The following example is from Blake (2001: 9):

(44) *Mīlitēs vident urbem*
'The troops see the city'

While verbs in a wide variety of languages exhibit tense marking, verbs in some languages do not. An example is Mandarin Chinese. The following is from Tiee (1986: 97):

	Zuótiān	*tā*	*xiě*	*zì*
(45)	yesterday	he	write	word
	'He wrote yesterday'			

(p. 193) Intersective variance can also be observed across languages insofar as a property might be exhibited by words in a different category from the category it is typically associated with in other languages. Tense marking is exhibited by verbs in a wide variety of languages. However, in certain languages it is also exhibited by nouns. In the following example from the Amazonian language Tariana (Nordlinger and Sadler 2004: 780), a marker of future tense is attached to the noun *unyane*:

	kayu-maka	*hi*	*waʃipeʃe*	*unyane-pena*
	SO-AFF	DEM:ANIM	Walipere	flood-fut
(46)	di-kakwa=pidana			
	3SG.NF-plan=REMPREP			
	'Thus Walipere was planning the future flood'			

Similarly, determiner marking is typically associated with nouns, but in certain languages, verbs may also be determiner marked. Croft (2001: 252) provides the following Palauan example:

	a	ngalęk	a	męnga	ęr	a	ngikęl
(47)	DET	child	DET	eat	OBJ	DET	fish
	'The child is eating the fish'						

Properties may also be cross-linguistically quirky insofar as they are exhibited by words in only one or very few languages. In a wide variety of languages, nouns are divided up into idiosyncratic semantic categories. In the Alaskan language Ahtna the noun class marker *d* denotes enclosed liquids and units of time, while the marker *n* denotes round objects, string-like objects, and liquids (Rice 2000: 326). Noun-class markers in the African language Kivunjo divide nouns up into sixteen semantically distinct subcategories, including body-parts, instruments, abstract qualities, and objects that come in pairs or clusters (Pinker 1994).

Such cross-linguistic variation makes it difficult to compare categories across languages; cross-linguistic generalizations about categories are also put in doubt. If a given category in one language exhibits a different set of properties from the corresponding category in a different language, how can we be sure that they are the same category and that it is valid to equate them? German adjectives exhibit a number of properties, such as case marking, that English adjectives do not. For Croft (2007: 417), this means that 'we have no syntactic basis for assuming that the English Adjective class is the same as the

German Adjective class'. Rather, he maintains: 'English Adjective and German Adjective are just language-specific categories, defined within each language, with no theoretical connection to each other.' While we may not be able to compare English and German adjectives in terms of properties that they do not have in common, there is still the possibility that we might compare them in terms of properties that they do have in common. One difficulty with this suggestion is that there are very few properties that a category of a given type is likely to exhibit across all languages. It (p. 194) would be wrong, for example, to assume that verbs across all languages exhibit tense marking, or that nouns in all languages exhibit determiner marking, as the above examples from Mandarin Chinese and Latin demonstrate. Nevertheless, it has often been argued that while there is no single grammatical property that a category will share across all languages, categories across languages might share common semantic properties. In particular, it has been maintained that in all languages nouns denote objects, verbs denote events, and adjectives denote properties (Croft 2001). However, it is clearly not the case that all nouns, verbs, and adjectives in all languages exhibit these semantic properties. There are many nouns in English that denote events rather than objects (e.g. *kiss, explosion, party*), and if we were to define categories in terms of semantic properties we would run the risk of misclassifying such nouns as verbs. Overall, then, it is difficult to determine what properties can be used to define lexical categories cross-linguistically.

There may also be cross-linguistic variation in the types of category exhibited by languages. For instance, adjectives and verbs can be clearly distinguished in English because they tend to exhibit a number of different properties. However, it is much harder to separate adjectives and verbs in, for example, the West African language Wolof because they tend to exhibit almost identical sets of properties. As a result, it is debatable whether such a language has a separate adjective category at all (McGlaughlin 2004). Similarly, in languages such as Mundari and Straits Salish, nouns, verbs, and adjectives exhibit such a similar set of properties that it can be argued that they form a single supercategory rather than three distinct categories (Bhat 1994). Moreover, cross-linguistic variation may arise not only because certain languages fail to exhibit common categories such as nouns, verbs, and adjectives, but also because they exhibit uncommon categories. A number of Australian languages, for instance, exhibit a category of words known as coverbs. Typically, a coverb does not inflect for tense, mood, or agreement but co-occurs with a finite verb that does. Consider the following example from Marra (Baker and Harvey 2010: 14).

(48)
rang=ng-anyi Ø-manuga
hit=1sg.S/3sg.O-TAKE.PC MA-rock
'I hit a rock'

Here, the coverb *rang* 'hit' combines with an inflected finite verb *nganyi* 'take'. If it were to be used without an accompanying coverb, *nganyi* would mean 'I was taking it'. However, when combined with a coverb such as *rang*, it functions as a light verb which conveys tense, aspect, mood, and agreement but little in the way of predicational meaning. Whereas verbs can occur independently of coverbs in Marra, coverbs cannot

occur without an accompanying verb. Moreover, as Baker and Harvey (2010) emphasize, verbs in Marra form a small, closed class with fewer than forty members, whereas coverbs form a large, open class. For all of these reasons, then, it is necessary to postulate a category of coverbs in languages such as Marra which is distinct from the category of verbs. Such differences in the categories that different languages exhibit, however, only (p. 195) exacerbate the problem of equating categories across languages. Should we equate the Marra verb category with the English verb category, or with the subcategory of light verbs such as *have* and *give* in English? Perhaps, instead, we should equate the English verb category with the set of Marra coverbs or even a combination of verbs and coverbs. Clearly, any such choice is liable to result in an imperfect match between the categories of Marra and English, since the two categories will inevitably exhibit radically different properties. It would be difficult to avoid Croft's (2007: 417) pessimistic conclusion that such categories 'are just language-specific categories, defined within each language, with no theoretical connection to each other'.

10.7 Conclusion

In contemporary linguistics, it is generally assumed that the grammatical behaviour of the words in a language can be adequately described in terms of a small set of lexical categories and the grammatical properties associated with them. Numerous studies, however, have shown that the behaviour of words is so complex and disorderly that it cannot be reduced to the grammatical properties associated with the lexical categories. As Taylor (2012: 280–81) has stated, 'The allocation of a word to one of the standardly recognized lexical categories, such as noun, verb, or adjective (or even to sub-categories of these, such as count noun, transitive verb, or predicative adjective) is grossly inadequate as a guide to the word's use in the language.' Nevertheless, many linguists have persisted in attempts to force words into lexical categories. Ultimately, all such attempts are inevitably arbitrary. Because there is never a good fit between lexical categories and the words that are forced into them, any choice of a set of lexical categories will be arbitrary. Similarly, this lack of fit also renders arbitrary any decision to assign a particular word to a given category or to associate a particular set of grammatical properties with a given category. In short, because lexical categories deny the complexity of word behaviour, any analysis of word behaviour that is based upon them is condemned to be arbitrary.

Mark Smith

Psychology, Open University

Mark Smith, Department of Psychology, Open University