

# Beyond Aristotle and gradience

## A reply to Aarts

William Croft

University of New Mexico

Aarts (2004) argues that the best way to model grammatical categories is a compromise preserving Aristotelian form classes with sharp boundaries on the one hand, and allowing gradience in terms of the number of syntactic properties that a category member possesses on the other. But the assumption of form classes causes serious theoretical and empirical problems. Constructions differ in their distributional patterns, but no a priori principles exist to decide which constructions should be used to define form classes. Grammatical categories must be defined relative to specific constructions; this is the position advocated in Radical Construction Grammar (Croft 2001). Constructionally defined categories may have sharp boundaries, but they do not divide words into form classes. Nevertheless, the most important traditional intuitions for parts of speech (Aarts' chief examples) are reinterpretable in terms of crosslinguistic universals that constrain distributional variation but do not impose Aristotelian form classes, gradable or not, on the grammars of particular languages.

### 1. Grammatical categories: Aarts' compromise

What is the nature of grammatical categories? The answers to this question have mirrored the answers to the more general question of categorization. For many centuries, the answer has been the Aristotelian one: categories have sharp boundaries, and are defined by individually necessary and jointly sufficient conditions.<sup>1</sup> In the last century, however, a new approach arose, partly based on philosophical considerations, partly on the basis of psychological experiments. In these experiments, gradient category behavior is consistently observed. This model goes under the name of 'prototype theory'. But the theoretical interpretation of gradient category behavior has been a matter of controversy. A common view is to take it as denying the existence of sharp boundaries. Boundaries

are fuzzy. A more extreme view is to conclude from this that category membership is a matter of degree.

Aarts (2004; henceforth MLG) proposes a model of grammatical categories that includes both gradience and Aristotelian grammatical categories with sharp boundaries. This model is intended to be a compromise between the traditional model, which denies gradience (or banishes it to the periphery), and the opposite extreme described above, which Aarts describes as the ‘gradience-is-everywhere’ view (MLG, 3). Aarts argues that categories do have sharp boundaries, but allows gradience in grammatical behavior within categories, such that some members display less of the grammatical behavior characteristic of the category than others. Aarts describes this as *subsective* gradience. For example, *utter* is a less central member of the English Adjective category than *thin*, because *thin* displays a wider range of grammatical behavior than *utter* (MLG:7):

- (1) a. a *thin* man (attributive position)
- b. he is *thin* (predicative position)
- c. very *thin* (intensified)
- d. *thin/thinner/thinnest* (graded)
  
- (2) a. an *utter* disgrace
- b. \*the problem is *utter*
- c. \*very *utter*
- d. \**utter/utterer/utterest*

In Aarts’ model, *utter* is an Adjective; there are no degrees of membership. But *utter* is not as good a member of the Adjective category as *thin* is.

Aarts also allows for less central members of a grammatical category to exhibit grammatical behavior of another category. He describes this as *intersective* gradience. Aarts argues that intersective gradience is much less common than subsective gradience, and a number of putative cases of it in the literature (to be discussed below) are in fact correctly analyzed as subsective gradience. An example that Aarts does conclude illustrates intersective gradience involves gerunds in various constructions, such as those in (3)–(4) (MLG, 32):

- (3) Brown’s deft *painting* of his daughter is a delight to watch.
- (4) Brown’s deftly *painting* his daughter is a delight to watch.

Aarts argues that *painting* in 3 is a noun that exhibits some verbal behavior, while *painting* in 4 is a verb that exhibits some nominal behavior.

I agree with Aarts that categories have both boundaries and some sort of internal structure not unlike gradience (see Croft and Cruse 2004, Chapter 4). But Aarts' model fails to address the most significant problems with both the 'gradience is everywhere' approach and the Aristotelian approach to grammatical categories. The shortcomings of Aarts' approach to categories and their boundaries are rarely remarked upon and in fact present a problem at least as pervasive in grammatical analysis as the problems associated with gradience in the prototype approach. Aarts is thus not alone. I address the problem here because Aarts' article argues for a general model of grammatical categories that is supposed to be adequate as 'systematized and idealized representations of the way we believe that grammar is mentally constituted' (MLG, 3). I argue in §2 that Aarts' model does not satisfy this theoretical goal. A model avoiding the problems of both gradience and Aristotelian categories, based on Croft (2001), is presented in §3, and applied to Aarts' examples in §4.

## 2. The Aristotelian model applied to grammatical categories

In §§3–5 of MLG (pp. 5–20), Aarts goes to some length to argue that gradience in syntactic behavior does indeed exist. I basically agree with Aarts' conclusion, or more precisely, I agree with Aarts that there is variation in the grammatical behavior among grammatical formatives<sup>2</sup> said to belong to the same category. In the formal syntactic tradition, this fact is frequently ignored or marginalized (but see Culicover 1999), and I hope that Aarts will succeed in bringing this fact to greater prominence in the formalist research tradition.

However, the audience of this journal is largely functionalist, and many members of this audience, generally more sympathetic to the 'gradience is everywhere' alternative, need to be persuaded that category boundaries, in particular sharp boundaries, are necessary for analysis. Unfortunately, Aarts presents no arguments as to why there must be grammatical categories with sharp boundaries. In Aarts' defense, I must point out that most grammarians assume that discrete form classes exist in every language and that it is just a matter of finding them, whether in English or in some less well documented language lacking an indigenous grammatical tradition. It is this assumption that is deeply problematic.

In many places, Aarts assumes not just that grammatical categories with sharp boundaries exist, but that the membership of certain formatives in their grammatical categories is obvious. He writes that both *thin* and *utter* (see (1) and (2) above) 'are *indisputably* adjectives' (MLG, 7; emphasis added). *Painting*

in examples (5) through (8) below ‘is *clearly* nominal’ (MLG, 18; emphasis added):

- (5) some *paintings* of Brown’s
- (6) Brown’s *paintings* of his daughters
- (7) The *painting* of Brown is as skilful as that of Gainsborough.
- (8) Brown’s deft *painting* of his daughter is a delight to watch. [= (3)]

The German form *sprechend+er/en* in (9) and (10) ‘has *clear* adjectival properties’ (MLG, 34; emphasis added):

- |      |       |         |           |                   |          |
|------|-------|---------|-----------|-------------------|----------|
| (9)  | ein   | mehrere | Sprachen  | sprechender       | Mann     |
|      | a.nom | several | languages | speaking:M.nom.sg | man.M.sg |
| (10) | einen | mehrere | Sprachen  | sprechenden       | Mann     |
|      | a.acc | several | languages | speaking:M.acc.sg | man.M.sg |

Yet these category membership judgements are not ‘clear’ and ‘indisputable’ to a ‘gradience-is-everywhere’ theorist, who needs to be persuaded that sharp category boundaries exist. Arguments must be presented for them.

Aarts does provide arguments for some of the cases he describes as ‘obvious’, using reasoning that is familiar to grammarians. In a footnote, Aarts states that ‘the fact that *utter* occurs in attributive position is a sufficient reason for assigning this word to the class of adjectives’ (MLG, 42, fn. 11). That is, attributive position is a sufficient condition for adjective category membership. *Painting* is ‘clearly nominal’ because it has a plural suffix in (5) and (6), a PP complement in (7) and a premodifying adjective in (8) (MLG, 18). That is, certain syntactic properties are again taken as a sufficient condition for category membership. The same is true of the German form in (9) and (10): its occurrence before a noun and its agreement in case and number are assumed to be sufficient conditions to define it as an adjective (MLG, 34). Aarts states that ‘*only* nouns can occur’ in the X role in the [Det Adj X] construction (MLG, 20, emphasis added), and that *design* in (11) ‘*crucially*...cannot be preceded by *the* or by an adjective’ (MLG, 30, emphasis added).

- (11) They intend to *design* new houses.

That is, occurrence in the X role in the [Det Adj X] construction is both a necessary and a sufficient condition for being in the noun category.

This is the normal mode of syntactic argumentation: find a distributional property that gives the category assignment you are looking for, or provides

the category boundary you are looking for. But there is a deep problem with this method of argumentation, which is hardly ever commented upon (Croft 2001, ch. 1). Aarts notes that ‘members of form classes do not possess the same distributional potential’ (MLG, 14). This is the reason that Aarts wants to incorporate gradience into his model of grammatical categories. In truth, this fact about language undermines the entire Aristotelian model of discrete form classes. Distributional potential is the occurrence of formatives (words, phrases, etc.) in different distributional contexts. But those distributional contexts are the very tests used to define the form classes (grammatical categories). These morphosyntactic tests (Croft 1991:6, cited in MLG, 30) are what are now called constructions in the generalized sense of that term in construction grammar (Fillmore, Kay and O’Connor 1988; Goldberg 1995; Croft and Cruse 2004). They include occurrence in the N role in the [Det Adj N] construction, or occurrence in the Adj role in the same construction, or inflection for the plural suffix, or agreement in case and number with a following element — the criteria taken by Aarts as ‘indisputable’, ‘clear’ and ‘crucial’ for determining the form classes of words.

Since distribution patterns differ by construction, different constructions define different boundaries for categories of words — sharp in most cases, but not the same, in fact often overlapping. The deep problem is that if formatives have different distributions in different constructions, how do we know which constructions should be used as tests for defining the form classes that are assumed by Aarts (and many others) to be the basis of grammatical description and analysis? The different constructions do not converge on the same boundaries, so this is a substantive empirical and theoretical problem. This problem was noted by both Bloomfield and Harris, who were among the first to pay close attention to methodological questions of syntactic analysis.

There is no principled answer to this problem. The standard solution, used by Aarts and most other grammarians, is to simply choose a subset of constructions, typically just one or two, as we have seen here. This is also the solution adopted by Bloomfield and Harris (Bloomfield 1933:269; Harris 1946:177) and their generative successors. In Croft (2001:30, 41) I call this form of argumentation *methodological opportunism*.

If one chooses certain constructions as criterial for defining form classes, then one must do something about the constructions that do not define form classes but define a different set of putative categories. Having assumed certain constructions to be criterial, Aarts treats the constructions with different distributions as being responsible for category gradience. For example, for Aarts, intensification and gradability are properties whose absence indicates a

less prototypical member of the adjective class (subsecutive gradience), while attributive position (the Adj role in [Det Adj N]) is criterial (but see below). However, no justification is given for these choices. From a theoretical point of view, it could just as well have been the other way around: gradability is criterial, and attributive position is contingent. This is not the usual decision, of course. But we want a theoretical basis for a decision as to what the form classes of English or any other language are, not a consensus view or simply an unquestioned grammatical tradition.

Aarts identifies this problem in his critique of another linguist, Jacobsson (1977). Jacobsson argues for intersective gradience between adverbs, prepositions and conjunctions in English.<sup>3</sup> The examples in question are in (12) and (13) (MLG, 16):

(12) I saw her *after* the concert.

(13) I saw her *after* she left the concert.

According to Aarts, Jacobsson would treat *after* as gradient between preposition and conjunction, because it has (for Jacobsson) a syntactic property criterial for prepositions — occurrence with an NP complement (example (12)) — and a syntactic property criterial for conjunctions — introducing a subordinate clause (13; MLG, 19). Aarts writes, ‘Jacobsson *aprioristically* assumes that these three distinct grammatical categories [adverb, preposition, conjunction] exist... What Jacobsson does not consider is the possibility that the taxonomy which he used as a starting point may in fact have been flawed’ (MLG, 18, emphasis in original). Aarts then notes that ‘it has been common to conflate the categories preposition and conjunction’. He gives the following examples (MLG, 19):

(14) a. John arrived *before* the last speech.  
 b. John arrived *before* the last speech ended.  
 c. John arrived *before* (hand).

(15) a. I haven’t seen him *since* the party began.  
 b. I haven’t seen him *since* the party.  
 c. I haven’t seen him *since*.

Aarts continues, ‘The *only* difference between the different instantiations of *before* and *since* is that in the a-sentences the preposition takes a nominal complement; in the b-sentences it takes a clausal complement, while in the c-sentences the preposition is intransitive’ (MLG, 19, emphasis added). If there is just one

category here, prepositions, then Jacobsson's examples reduce to subjective gradience (MLG, 19–20).

Aarts does not consider the possibility that his proposed alternative taxonomy is equally flawed. No reason is given for ignoring the difference in complements. Aarts downplays the distributional difference by using the word *only* in the quoted passage. In a footnote, Aarts states, 'the complement-taking properties of these elements [*before* and *since*]...create subcategories within form classes, hence *subcategorization*, but not new form classes' (MLG, 41, fn. 7, emphasis in original). Why? Although Aarts' view follows recent tradition, this is a stipulation just as much as Jacobsson's decision to keep the categories apart by virtue of the types of complements (if any) the formative takes. The dispute between Aarts and Jacobsson is reflected in many other debates on grammatical categories in the formal and typological syntactic literature: one side lumps formatives together into a single category, while the other splits them. 'Yet there is no a priori way to resolve the question: the "lumper" overlooks the mismatches in distribution, and the "splitter" overlooks the generalizations' (Croft 2001:32).

Aarts also takes the lumper's side by adopting Hudson's (2000) argument that *that* in (16) as well as (17) is a pronoun:

(16) *that* house

(17) What is *that*?

Aarts argues that *that* in (16) and (17) are pronouns 'which may, or may not, take a complement' (MLG, 29). Again, this distributional difference is downplayed. This analysis would require that Nouns (or N-bars) such as *house* in (16) are complements. But they do not have the same distribution as other complements:

(18) \*I bought house. [Verb + "Complement"]

(19) \*that John's house [*that* + Complement]

Once again, the splitters could point to the evidence in (18) and (19), while the lumpers will point to the evidence in (16)–(17); and again, there is no a priori way to choose between the two.

Aarts offers one explicit reason and one implicit reason to choose some syntactic properties as criterial and to ignore others. The explicit reason is that the properties are unique to the members of the form class (MLG, 32). But uniqueness fails for one of the criteria that Aarts uses the most, the prenominal attributive position used to define English Adjectives (see above). In footnote

11, Aarts asserts that prenominal position is sufficient to make *utter* an adjective (quoted above). But he then notes that nouns can also occur prenominally; that is, it is not a unique criterion. In footnote 20 Aarts allows words that cannot occur attributively such as *afraid* to be adjectives, so even attributive position is not a necessary condition for adjectivehood. Finally, prenominal position is also found for the “transitive pronoun” *that* in Aarts’ own analysis (see (16)). Prenominal attributive position is an important construction for defining adjectives (see §3), but not for the reason that Aarts gives.

The implicit reason in Aart’s article for ignoring distributional differences is that the forms in the two different distributional contexts are identical. Aarts writes of Hudson’s analysis of *that* in (16) and (17), ‘We could also say that in this particular case the grammar is not trying to delude us, and that the *thats* in [(16)] and [(17)] do not only look the same, they *are* in fact the same’ (MLG, 29, emphasis in original). This is also Aarts’ reason for ignoring the distributional differences between adverbs, prepositions and conjunctions: formally identical words such as *before* and *since* in (14) and (15) would otherwise be in two separate categories (or require intersective gradience; *ibid.*).

If this reasoning were taken to its logical conclusion, it would lead to erasing the distinction between Noun and Verb in English, for example:

- (20) a. She *sliced* the sausage.  
 b. I’ll take two *slices*, please.
- (21) a. This *pipe* is six inches thick.  
 b. They will *pipe* natural gas across Daghestan.

In (20) and (21), we could either posit two *slice*’s and two *pipe*’s, one a Noun and the other a Verb. Or we could follow Aarts’ implicit reasoning, and say that there is a single category Verb, and many Verbs also occur in referential constructions. Most linguists, myself included, would reject this reasoning (though it has been proposed for a number of languages, and I have heard it proposed even for English); see §3. But it shows that analyzing identical forms as belonging to a single category is not a sufficient reason for ignoring distributional differences (Croft 2001:76).<sup>4</sup>

Finally, the problems of methodological opportunism are multiplied exponentially as soon as the principle is applied to more than one language. Aarts gives one example from another language, German (see (9)–(10) above). Even though German is genetically, areally and typologically close to English, the example suffices to illustrate the problem. As noted above, Aarts states that *sprechend+er/en* is ‘clearly’ adjectival because it occurs in prenominal position



and agrees with the following noun in case and number (its ‘Verbal’ property of taking a direct object is an instance of intersective gradience). If we return to English and ask, which English words fit the criteria for Adjectives in German, the closest match is not *thin* or *utter*, but *this* and *that*, since they agree in number (though not case) with the following noun:

(22) *this* box/*these* boxes

(23) *that* chair/*those* chairs

These distributional facts give us two options for analysis. One can say that English has only two adjectives, *this* and *that*. If we take this option, we still have no basis for saying why agreement with the following noun is the criterion for adjective class membership, rather than some other criterion — the point I have been arguing so far. Or one can say that there is an English Adjective class, which includes *thin* and *utter* but not *this* or *that*. If we take the latter option, then we have no syntactic basis for assuming that the English Adjective class is the same as the German Adjective class. Hence, we have no basis for saying that there is a category ‘adjective’ in grammatical theory that is valid across languages. English Adjective and German Adjective are just language-specific categories, defined within each language, with no theoretical connection to each other. I do not think Aarts intends to take this position.

This is of course the problem of parts of speech, or syntactic categories in general, found in studies of language that take seriously comparative data from more than one language (see Croft 2001, Chapter 2). For example, in many languages, predication of an “adjective” is syntactically like predicating a “verb”: there is no copula, and the predicated “adjective” agrees with the subject. Some linguists argue that there are no adjectives in such a language: since the “adjective” agrees with the subject and there is no copula, it is merely a verb. There is no justification for why these criteria should be the absolute criteria for defining adjectives. Other linguists find other constructions that distinguish “adjectives” from “verbs”, and argue that such languages do distinguish “adjectives” from “verbs”. But there is no syntactic basis for saying that the two categories are the same categories as Adjective and Verb in English, since different constructions are used to define the categories. So the latter set of linguists have not demonstrated that all languages (or the language in question) have “adjectives” as well as “verbs”. Instead, they have simply shown that the language in question makes a grammatical distinction between two sets of formatives. And the pro-adjective and anti-adjective linguists will never agree, because there is no a priori basis for choosing which criteria are the “right” ones, those used by

the pro-adjective camp or those used by the anti-adjective camp, or what to do when criteria conflict.

In sum, there are serious, in fact insurmountable, problems in positing Aristotelian categories in grammar. There is great variation in distributional patterns for formatives. Hence the number of word classes in one's analysis depends crucially on which constructions are chosen as tests for belonging to word classes. Aarts chooses a particular construction as criterial for a word class, and the mismatches of other constructional distributions are dealt with by subsective and intersective gradience. But there are no *a priori* criteria for choosing one construction over another as the defining test for a word class. Aarts' choices are essentially arbitrary. The problem gets worse across languages. There is even more variation, hence even more difficulties in applying the same criteria across languages. And if one uses different constructions in different languages, then one has no basis for asserting that the categories found in different languages are instances of one and the same grammatical category valid across languages, and hence valid for syntactic theory.

Why does Aarts persist in assuming that there are Aristotelian categories in grammar in the face of these problems? Aarts writes, 'a certain degree of idealization is necessary in order for a description of a language to be possible at all, so as to make sense of the wealth of linguistic facts that we face within particular languages and crosslinguistically' (MLG, 37). But the idealization of Aristotelian categories is incompatible with the huge amount of variation in distributional facts. Adding gradience is just a patch required for a flawed model of grammatical categories.

Nevertheless, there are intuitions that virtually all linguists share, including Aarts and myself, that have guided some of the choices Aarts and his predecessors have made: the attributive construction is significant for defining "adjective", and heading a noun phrase containing a definiteness marker such as *the* is significant for defining "noun". Is it possible to capture these intuitions, which I share with Aarts, without falling into methodological opportunism and without assuming grammatical categories with sharp boundaries of the sort that Aarts wishes to defend? The next section argues that it is.

### 3. A radical constructional approach to grammatical categories

Virtually all ordinary working linguists, whether analyzing English or some less documented or undocumented language, assume that the only way to describe the language is to identify a set of word classes with boundaries (sharp or

fuzzy). Is there really any way to look at grammar except as a set of formatives divided into a small set of discrete form classes defined by certain constructions? Fortunately, the answer to this question is ‘yes’.

Let us return to one of the chief criteria used by Aarts for defining the English Adjective category, prenominal position. Aarts considers prenominal position to be a sufficient condition for categorizing *utter* in the same way as *thin*, and concedes that Nouns can also occur in this constructional context (MLG 42, fn. 20):

(24) an *utter* disgrace [= (2a)]

(25) a *thin* man [= (1a)]

(26) a *jewelry* box

Aarts assumes that this construction tells us something about the inherent, fixed word class membership of the words in examples (24)–(26). But the facts observed in examples (24)–(26) do not tell us anything about inherent word class membership. They actually tell us something *about the attributive construction*. Likewise, the facts in (27)–(29) tell us something about the intensification construction, not about the relative centrality of *thin*, *utter* and *false* as “adjectives”:

(27) very thin [= (1c)]

(28) \*very *utter* [= (2c)]

(29) \*very *false*

The categories are defined by the construction (the morphosyntactic ‘test’ or ‘criterion’), not the other way around. This is what distributional analysis *is*: namely, identifying the occurrence of certain formatives in certain constructions and not others. The central point is that the specific constructions used to define grammatical categories *matter* in grammatical analysis. The same is true of gradience models of grammatical categories, including Aarts’ model. The problem with ‘gradience is everywhere’, and also with Aarts’ incorporation of gradience into his theory, is that they are purely quantitative. They simply count the number of constructions characteristic of that category’s behavior that the word occurs in, without considering which constructions are being counted.

The real problem with both the quantitative approach to grammatical gradience and the Aristotelian model of grammatical categories is that neither pays attention to the *actual constructions* used to justify category membership

(for Aristotelian categories) or goodness of membership (for gradience). In the quantitative model of grammatical gradience, all that matters is how many constructions the formative is acceptable in, not which ones they are, or why. In the Aristotelian model of grammatical categories, all that matters is that there is a construction that supports the existence of the category, not which construction it is, or why.

The real work for the grammarian, and the genuine key to the nature of grammar, lies in describing the actual relationship between formatives and the constructions they occur in, and in understanding why they vary in the way that they do. The correct view of grammatical categories is not ‘gradience-is-everywhere’ but ‘variation-is-everywhere’: variation in distribution patterns of different formatives in the same and different constructions within a language, and variation in distribution patterns of equivalent formatives and equivalent constructions across languages. When Langacker (1987:19) speaks of ‘rigorous description’ (quoted in MLG, 19), this is what he has in mind.

A model of what a speaker actually knows about the grammar of her language has to specify which syntactic properties each formative has, that is, which constructions each formative occurs in (or not). But once you have that, you don’t need gradience — it is epiphenomenal. The same is true of Aristotelian categories. A model of what a speaker actually knows about the grammar of her language has to specify which constructions each formative occurs in (or not). Once you have that, you don’t need Aristotelian categories. Any Aristotelian categories you posit, either on the basis of the usual constructions invoked or on the basis of any other construction you choose, are merely epiphenomenal.

This is the approach advocated in Radical Construction Grammar (Croft 1999, 2001, 2004). In Radical Construction Grammar, the constructions used to define categories are simply that: they are the basic units of grammatical representation, and they define the categories of formatives which occur in the construction.<sup>5</sup> Attention is therefore directed to the complex and variable distributional relationship between particular constructions and the formatives that occur in specific roles in the constructions, both within and across languages.

#### 4. Applying a radical constructional analysis to Aarts’ examples

How would a radical constructional analysis of the examples in Aarts’ paper look? The Radical Construction Grammar analysis of parts of speech does not

have Aristotelian grammatical categories of the sort envisioned by Aarts for particular language grammars. There are categories for each construction and each constructional role in a language. These construction-specific categories will have sharp boundaries to the extent that there are sharp acceptability judgements of what can and cannot occur in the relevant constructional role. In this sense, the categories are Aristotelian. But they do not lead to a small set of mutually exclusive word classes, which is what Aarts assumes we must posit. Instead, there are overlapping categories of formatives representing their diverse distributional behavior — which is what a speaker actually knows about her language. The real question about the nature of grammar is, then, is there any systematicity to the distributional diversity of formatives, given that they do *not* fall into large, discrete form classes? And does this systematicity have anything to do with the traditional categories of ‘noun’, ‘verb’ and ‘adjective’ that figure prominently in Aarts’ article, and the intuitions that Aarts, I and many others share about these parts of speech? Again, the answers to these questions are ‘yes’.

The primary construction that Aarts uses for defining the category of Adjective in English (and German) is the attributive construction, and the primary construction Aarts uses for defining a Noun in English is the referring construction. More precisely, Aarts uses the attributive role Adj in the [Det Adj N] construction for defining adjectives, and the referring expression role N in the same construction for defining nouns. In this respect, Aarts is following tradition. There are in fact good empirical linguistic reasons for doing so. The referring construction and the attributive construction, along with the predication construction, have the function of symbolizing the propositional acts of reference, modification and predication, respectively. An analysis of these functions can be found in Croft (1991:101–26). Essentially, reference and predication correspond to Searle’s propositional acts (Searle 1969), one of the levels of speech acts (Austin 1962; Clark 1996) or information structure in an utterance. Modification performs a secondary propositional act function (see Wierzbicka 1986; Croft 1991:122–23). By virtue of their related information-structure functions, these constructions form a paradigmatic set and can be set off from other constructions that perform other functions.

In invoking the functions of the relevant constructions, we go beyond the ‘purely morphosyntactic’ criteria of distributional analysis, or so it would seem. This is in fact not true. We are still concerned with which formatives occur in which constructions. What we are trying to do is identify a subset of constructions and formatives which may share consistent patterns of distributional behavior, and find an explanation for those patterns. The criteria for identifying

those constructions and formatives are symbolic, that is, what meaning or function the constructions and formatives encode. The use of functional criteria is also necessary for identifying equivalent formatives and constructions across languages (Croft 2003:13–19). Prenominal position in German is equivalent to prenominal position in English because they both perform the attributive (modification) function. The functional equivalence remains even though the German construction has agreement in case and number with the head and the English construction does not. Functional equivalence allows us to develop a crosslinguistically valid and empirically supportable theory of parts of speech that a ‘purely’ morphosyntactic approach cannot. Functional equivalence also allows us to provide a theoretical basis for the pretheoretic intuitions that have guided Aarts and other linguists to using the attributive, predication and referring constructions for analyzing parts of speech.

Functional equivalence also allows us to control for function in a way that a ‘purely’ morphosyntactic account cannot. The examples of ‘noun’-‘verb’ equivalence in English given in (20) and (21) in §2 did not take into consideration the fact that there is a substantial semantic difference between the two occurrences of *slice* and *pipe* (which is moreover largely idiosyncratic). In a ‘purely’ morphosyntactic analysis, we cannot take this fact into consideration, or differentiate it from the very small semantic difference between *a thin man* (example (1a)) and *he is thin* (example (1b)). As a result, we cannot capture significant generalizations about the relationship between meaning, form and distribution in a ‘purely’ morphosyntactic account (see Croft 2001:65–75).

Our functionally-based reasoning has also provided a justification for ignoring differences in complement-taking properties when analyzing parts of speech, a point on which I criticized Aarts in §2. To see this, we must ask what it means to say that some formatives take complements of certain types and others do not. It means we are describing the distribution of formatives in a variety of argument structure constructions. This is of course an interesting research question in its own right, but it involves a different set of constructional functions than the propositional act functions. Of course, these might turn out to interact with propositional act functions in interesting ways. But this is an empirical question to be answered, not something to be excluded aprioristically by asserting that complement-taking properties have nothing to do with parts of speech.<sup>6</sup>

Pursuing this line of reasoning a bit further, it also suggests that lumping together “prepositions” and (subordinating) “conjunctions” (examples (12)–(15)) is also theoretically unfruitful. Argument structure (the functional domain of “prepositions”) and clause linkage (the domain of “conjunctions”) are

functionally quite distinct. Nevertheless, there are also semantic similarities in terms of a figure-ground construal of the head/main clause and the complement/subordinate clause (Talmy 1978; Croft 2001:329–35). “Adpositions” and “subordinators” are also grammatically related diachronically in that a well-documented grammaticalization path leads from adpositions to subordinators (see Meillet 1915/1921 and Lehmann 1982/1995:67 for Romance, and Genetti 1986, 1991 for the Bodic subgroup of Sino-Tibetan). One effect of grammaticalization paths is that because the grammaticalization process is gradual, in many languages the forms in question share some properties with “true adpositions” and some properties with “true conjunctions”. The properties shared differ from language to language (see Croft 1991:142–46). An Aristotelian approach requires us either to lump “adpositions” and “subordinators” together (as Aarts does), or to treat them as completely independent (as Jacobsson does). Either decision prevents us from identifying and discovering the universal diachronic generalization and the language-specific synchronic facts that follow from it.

I now return to the propositional act constructions and the distribution of formatives that are found in them, using Aarts’ examples. Different semantic classes of formatives can occur in different propositional act functions. For example, property words can be predicated, as in (30):

(30) He is *thin* [= (1b)]

Actions can be referred to, as in (31) or (32):

(31) Brown’s deft *painting* of his daughter is a delight to watch. [= (3)]

(32) Brown’s deftly *painting* his daughter is a delight to watch. [= (4)]

Object words can function as modifiers, as in (33):

(33) a *jewelry* box [= (26)]

However, the relationships between semantic class of formative and role in different propositional act constructions is systematically asymmetric in a cross-linguistic perspective. In particular, the following combinations of semantic class and propositional act function have a special status:<sup>7</sup>

- (34) a. reference to an object  
 b. modification by a property  
 c. predication of an action

These three combinations of semantic class and propositional act will be grammatically encoded by no more morphemes than the combinations not listed in (34) (the structural coding criterion of typological unmarkedness; Croft 2003, Chapter 4), and these three combinations will possess at least as many inflectional distinctions as the combinations not listed in (34) (the behavioral potential criterion of typological unmarkedness; *ibid.*). These combinations are typological prototypes (Croft 2003, Chapter 6). Typological prototypes are *not* language-specific categories. Instead, they are generalizations about the grammatical expression of semantic and functional categories that constrain the variation in distributional patterns across languages.

What has happened to parts of speech as Aristotelian form classes? The typologically unmarked combinations of semantic class of formative and propositional act function of the relevant constructions in (34a-c) are related to previous definitions of noun, adjective and verb as word classes: the semantic class corresponds to the traditional notional definition, and the propositional act is the function of the constructions most commonly used for defining parts of speech as word classes in the structuralist/generative tradition. I am happy to label these privileged pairings with the terms ‘noun’, ‘adjective’ and ‘verb’ (Croft 2001:89). The problem lies with interpreting the combinations in (34) as claims about Aristotelian form classes in specific languages such as English. This inference is incorrect. We cannot even use (34) as a means to identify a particular construction or set of constructions for defining parts of speech as Aristotelian form classes, because the constructions relevant for the theory of parts of speech overlap in various ways.

We may now analyze further examples in Aarts’ article in this model. Aarts discusses differences in behavior for three English “adjectives”, *thin*, *utter* and *afraid*: *utter* does not occur in degree constructions (examples (1c)–(1d)), and cannot be predicated (example (1b)), while *afraid* can only be predicated. These behavioral differences conform to the general principles given above. *Thin* is a property as narrowly defined; in fact, it is part of Dixon’s (1977) core semantic “adjective” class. Thus we would expect it to occur in the attributive construction, and to occur in degree constructions; the expression of degree represents part of the behavioral potential of modifiers (Croft 1991:79). *Utter* and *afraid* are not property words. They therefore may turn out to lack the behavioral potential of property words (as with *utter*), and even not to occur in the attributive construction (*afraid*). There are sharp boundaries in the language-specific categories, defined by the occurrence in the English constructions encoding modification, predication and degree. But the boundaries overlap and no particular construction’s distributional boundary is privileged. There is also



gradience in this model, but the gradience has to do with the relationship between lexical semantic classes and the propositional act functions in (34), not with specific form classes.

The structural coding criterion states only that the typologically unmarked combination of semantic class and propositional act function will be coded by *no more morphemes* than other combinations. Hence, the fact that object words in English can occur in the same attributive construction as property words (MLG, footnote 11; example (33)) is not a counterexample to the typological universal. A genuine counterexample would be a language in which a property word in an attributive construction required overt coding of the modification function, while an object word did not; or an action word in a predication construction that required a “copula” while a predicated property word did not. In fact, these typological universals for semantic classes of formatives and propositional act constructions are extremely robust.

An important reason for not positing sharp boundaries between word classes is that doing so would miss language universals that cut across the boundaries, applying both to subclasses within an Aristotelian category (Aarts’ subsective gradience) and across word classes (Aart’s intersective gradience). For example, Aarts refers to the gradability of English Adjectives (MLG, 7). There are actually three gradability constructions in English: (i) a pair of suppletive gradability sets, *good/better/best* and *bad/worse/worst*; (ii) the inflectional constructions with *-Ø/-er/-est* (e.g. *thin/thinner/thinnest*); and (iii) a periphrastic set of gradability constructions, *Ø Adj/more Adj/most Adj* (e.g. *intelligent/more intelligent/most intelligent*). The inflectional set has been losing members over time and is now restricted to a set of monomorphemic adjectives with certain prosodic properties.

These three constructions define three grammatical categories: one including *good* and *bad*, one including the set of inflectionally gradable “adjectives”, and one including the remaining “adjectives”. But the three categories can be ranked in terms of unmarkedness of behavioral potential, with the suppletive category least marked and the periphrastic most marked. Category membership cannot be predicted, of course, but it is motivated by a well-documented ranking of property concepts in terms of markedness as adjectives in the sense of (34b) (Dixon 1977; Stassen 1997:168–69). ‘Good’ and ‘bad’ — value expressions — belong to the most core adjectival property concepts. The English words denoting other core adjectival concepts — dimension (*big, small, wide, narrow, thick, thin, tall, short, high, low*), age (*old, young, ripe*) and speed (*fast, quick, slow*) — are all in the inflectional gradability category. Thus, the “subclassification” of “adjectives” (subsective gradience) determined by the gradability

constructions of English actually conforms to the typological universals of parts of speech summarized in (34). If we simply called them all “adjectives”, by interpreting the generalizations in (34) as defining Aristotelian form classes, we would miss this generalization.

The parts-of-speech universals cut across Aristotelian word classes, motivating the behavior of categories representing the non-privileged combinations not in (34), such as gerunds (reference to action) and participles (actions used as modifiers, as in the German *sprechend+er/en* examples in (9) and (10)). Aarts is obliged to decide for each type of gerund construction whether the gerund form is really a noun or really a verb, since he is committed to an Aristotelian grammatical category model. Aarts writes:

Martin Haspelmath asks (p.c.) “Suppose there are two categories with 50 properties each. If an observed item has 23 or 24 properties of one of them, does that make it so different from another item that has exactly 25?” I would answer ‘yes’ to this question. One or two properties can make all the difference, as a comparison between *Brown’s painting of his daughter* and *Brown’s painting his daughter* makes clear: the first of these is nominal, while the second is verbal. (MLG, 38)

Aarts’ conclusion follows from his Aristotelian assumption; yet Haspelmath is questioning the utility of this very assumption. But if one discards the Aristotelian assumption, and recognizes that both constructions and formatives have functions, then the intermediate status of gerunds (and participles) is naturally explained: a gerund is a formative that is semantically closer to the ‘verb’ combination (example (34c)) but is used in a propositional act role that is characteristic of the ‘noun’ combination (example (34a)). Hence it is not surprising that gerunds display some properties of “nouns” and some of “verbs”, and in fact in different mixtures.

Moreover, there are implicational scales governing which properties of expressions referring to actions will be “nounlike” and which will be “verblike” (Comrie 1976; Croft 1991:83–85, 1995:82, 2001:355–57; Koptjevskaja-Tamm 1993:257; Cristofaro 2003, ch. 10). The implicational scales constrain the possible combinations of “nominal” and “verbal” properties of gerunds. For example, example (31) uses a “nominal” encoding of both the subject and object arguments of *painting*, while example (32) uses a “nominal” encoding of the subject and a “verbal” encoding of the object. The universals predict correctly the absence of a gerund construction with a “verbal” encoding of the subject and a “nominal” encoding of the object. As with the adposition-subordinator grammaticalization path, if we draw a sharp line between “nominal” and

“verbal” gerunds, then we miss the generalizations linking the different constructions used in referring to actions. And if we simply count the number of “nominal” and “verbal” properties, we miss the implicational universal constraining the coding of the subject and object arguments.

## 5. Conclusion

In this paper, I have argued that one must take seriously the actual constructions used in analyzing the grammatical categories of a particular language. The facts of distributional variation show that constructions do not define categories that converge on a small set of mutually exclusive form classes (word classes). Aarts, like most other linguists, assumes that there must be word classes, and chooses particular constructions to justify those word classes; but there is no theoretical basis for doing so. Many other linguists discount or even ignore the conflicting evidence from other constructions. Aarts progresses beyond that strategy by adding a layer of category gradience based on the evidence from other constructions. But there is still no basis for privileging some constructions as establishing form classes, while demoting other constructions to signalling gradience within the form classes.

But grammatical analysis does not have to be this way. The grammatical knowledge of a speaker must include all distributional facts. If so, then we do not need to posit Aristotelian form classes. This is the approach advocated in the Radical Construction Grammar model (Croft 2001). In this model, categories are defined by the constructions of the language — as the distributional method dictates. These categories have boundaries which are more or less sharp.<sup>8</sup> None of the constructions has a privileged status in defining word classes.

This is not to say that ‘everything goes’. The starting point is always a detailed distributional analysis, which describes what a speaker knows. But this is not saying, ‘there is no pattern’. There *are* patterns which are empirically verifiable and valid generalizations across languages. In order to identify such universals, we must classify the constructions, and the formatives whose distribution they define, in terms of their functions. This classification also allows us to compare equivalent constructions and formatives across languages. In the case of parts of speech, the focus of Aarts’ examples, there are universals of how semantic classes of formatives are distributed in language-specific constructions encoding propositional act functions, such as the typological prototypes in (34) and the ranking of constructions for gerund-like constructions referred

to in §4. These universals constrain variation in the grammatical categories defined by language-specific constructions that perform the relevant propositional act functions. That is, these universals constrain a speaker's knowledge and acquisition of grammar. I believe that this structure represents the way that grammar is mentally constituted.

## Notes

1. An anonymous referee has suggested that a better description would be 'Donatine', since Aristotle did not devote much attention to grammatical categories; but I will follow Aarts' convention here.
2. Aarts uses the term 'formative' to describe any grammatical element, from a morpheme to a word to a larger syntactic unit.
3. Aarts gives no examples from Jacobsson's paper, so my comments are based on Aarts' application of Jacobsson's method of analysis to Aarts' own examples.
4. In discussing gerunds and the example of it ("noun" or "verb"), Aarts proposes using as syntactic properties only those that can occur together in the same construction (MLG, 30). But Aarts does not follow his own proposal when he lumps together prepositions and subordinating conjunctions, since a word such as it cannot both take a nominal complement and introduce a subordinate clause in the same construction.
5. The term 'radical' refers to the thesis that complex units, namely constructions, are basic, that is, not defined in terms of the types of the smaller units the construction is made up of. This is a nonreductionist model of grammar, just as the Gestalt theory of perception is a nonreductionist theory of perception. Instead, a construction is a whole with a distinctive formal structure and a distinctive semantic/discourse function, both of which differentiate it from other constructions. See Croft (2001:47–61) for further discussion.
6. In fact, there is an interaction of argument structure with the propositional act constructions of predication vs. reference, that is, action predication vs. action nominalizations (Comrie 1976; Croft 1991:84–85, 2001:354–61; Koptjevskaja-Tamm 1993; Cristofaro 2003, ch. 10). But this interaction has to do with the argument roles of the predication construction, not the predicate role, which is the role relevant to parts-of-speech theory.
7. The definitions of the semantic classes are as follows: object = nonrelational, stative, permanent, nongradable; property = relational, stative, permanent, gradable; action = relational, dynamic, transitory, nongradable. For discussion of the semantic properties see Croft (1991:62–65, 2001:87).
8. Category boundaries may not be sharp if acceptability judgements are not categorical, or if they vary across speakers or occasions — as they certainly do.

## References

- Aarts, Bas. 2004. Modelling linguistic gradience. *Studies in Language* 28: 1–49.
- Austin, J. L. 1962. *How to do things with words*. Cambridge MA: Harvard University Press.
- Bloomfield, Leonard. 1933. *Language*. New York NY: Holt, Rinehart and Winston.
- Clark, Herbert H. 1996. *Using language*. Cambridge: CUP.
- Comrie, Bernard. 1976. The syntax of action nominals: A cross-language study. *Lingua* 40: 177–201.
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: OUP.
- Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago IL: University of Chicago Press.
- Croft, William. 1995. Review of Maria Koptjevskaja-Tamm, *Nominalizations*. *Nordic Journal of Linguistics* 18:75–83.
- Croft, William. 1999. Some contributions of typology to cognitive linguistics (and vice versa). *Cognitive linguistics: Foundations, scope and methodology*, Theo Janssen and Gisela Redeker (eds.), 61–93. Berlin: Mouton de Gruyter.
- Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: OUP.
- Croft, William. 2003. *Typology and universals*, 2nd edition. Cambridge: CUP.
- Croft, William. 2004. Logical and typological arguments for Radical Construction Grammar. *Construction Grammar(s): Cognitive and cross-language dimensions*, Mirjam Fried and Jan-Ola Östeman (eds.), 273–314. Amsterdam: John Benjamins.
- Croft, William & Cruse, D. Alan. 2004. *Cognitive linguistics*. Cambridge: CUP.
- Culicover, Peter. 1999. *Syntactic nuts: hard cases, syntactic theory, and language acquisition*. Oxford: OUP.
- Dixon, R.M.W. 1977. Where have all the adjectives gone? *Studies in Language* 1:19–80.
- Fillmore, Charles J.; Paul Kay and Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64: 501–538.
- Genetti, Carol. 1986. The development of subordinators from postpositions in Bodic languages. *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, Vassiliki Nikiforidou; Mary VanClay; Mary Niepokuj and Deborah Feder (eds.), 387–400. Berkeley CA: Berkeley Linguistics Society.
- Genetti, Carol. 1991. From postposition to subordinator in Newari. *Approaches to grammaticalization*, Vol. 2, Elizabeth Closs Traugott and Bernd Heine (eds.), 227–255. Amsterdam: John Benjamins.
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago IL: University of Chicago Press.
- Harris, Zellig S. 1946. From morpheme to utterance. *Language* 22: 161–183.
- Hudson, Richard A. 2000. Grammar without functional categories. *The nature and function of syntactic categories* [Syntax and Semantics 32], Robert D. Borsley (ed.), 7–36. New York NY: Academic Press.
- Jacobsson, Bengt. 1977. Adverbs, prepositions and conjunctions in English: A study in gradience. *Studia Linguistica* 31: 38–64.
- Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London: Routledge.

- Langacker, Ronald W. 1987. *Foundations of cognitive grammar*, Vol I: *Theoretical prerequisites*. Stanford CA: Stanford University Press.
- Lehmann, Christian. 1982/1995. *Thoughts on grammaticalization: A programmatic sketch*, Vol. I. Cologne: Institut für Sprachwissenschaft. [Arbeiten des Kölner Universalien-Projekts 48]. (Reprinted by LINCUM Europa, Munich, 1995).
- Meillet, Antoine. 1915/1921. Le renouvellement des conjonctions. *Linguistique historique et linguistique générale*, 159–174. Paris: Klincksieck. (Originally published in *Annuaire de l'École Pratique des Hautes Études*, Section historique et philologique, 1915).
- Searle, John R. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge: CUP.
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: OUP.
- Talmy, Leonard. 1978. Figure and ground in complex sentences. *Universals of human language*, Vol. 4: *Syntax*, Joseph H. Greenberg; Charles A. Ferguson and Edith A. Moravcsik (eds.), 625–652. Stanford CA: Stanford University Press.
- Wierzbicka, Anna. 1986. What's in a noun? (or: how do nouns differ in meaning from adjectives?). [\*Studies in Language\* 10: 353–389.](#)

*Author's address:*

Prof. Dr. William Croft  
Dept. of Linguistics  
Humanities Bldg 526  
University of New Mexico  
ALBUQUERQUE NM 87131-1196  
USA  
wcroft@unm.edu