

Chapter 2

Background

The focus of this chapter is to explain the concept of lexical flexibility, consider its criticisms, and offer a more robust, functionally-grounded definition instead. I first briefly describe how flexible approaches to lexical categories developed as a response to weaknesses in traditional theories of parts of speech. I then survey the landmark studies and important findings on lexical flexibility, along with criticisms of this research. Following that, I present the typological markedness theory of lexical categories, which states that lexical categories are merely emergent markedness patterns regarding how different semantic classes of words are used for different discourse functions. I conclude by offering a revised formulation of lexical flexibility which is in line with typological markedness theory.

2.1 Introduction: Approaches to lexical flexibility

The field of linguistics as a whole, and the subfield of typology in particular, is undergoing a radical shift in how we understand lexical categories, along primarily two dimensions. The first dimension is our understanding of what lexical categories are a property *of*. Early researchers viewed categories as universal properties of both language and languages (CITE: Haspelmath on g-language vs. p-language; add these terms in parentheses). I call this the *universalist* position. After Boas, many researchers then came to view categories as language-specific, with patterned similarities across languages. I call this the *relativist* approach. Most recently, some researchers view categories as typological patterns rather than properties of

any particular language. This is the *typological* position, and the one I adopt here.

The second dimension of historical change in linguistic theories of categories is in the *nature* of the categories themselves. In the Classical tradition, categories were thought to be categorical and well-defined by a set of necessary and sufficient conditions (in the tradition of Aristotle). After the cognitive turn in the 1960s and 1970s, many linguists came to view categories as prototypical, with some members of a category being more central, or better exemplars, than others. Cognitive research into the nature of idioms then led to the development of construction grammar, which sees language as consisting of a network of constructions rather than monolithic categories. I adopt a constructional approach to categories in this thesis.

These theoretical paradigm shifts are summarized in (10). At each stage of development, there has not been a wholesale displacement of previous theories (CITE: Kuhn). There are still many who regard word classes as universal and categorical, and the typological-constructional approach is still nascent.

- (10) a. universal > language-specific > typological
b. categorical > prototypical > constructional

Section 2.2 gives a synopsis of these theoretical positions and shows how research on lexical flexibility developed in recognition of the shortcomings of traditional approaches. Section 2.3 summarizes the key concepts and findings that have arisen from the research on lexical flexibility. Such research, however, is not without its own shortcomings. Section 2.3 also presents the main criticisms that have been leveled against flexible analyses of word classes. Section 2.4 then presents an alternate, functionally-oriented approach—the typological-constructional perspective. The final section of this chapter (§2.5) then applies this functional perspective to formulate an improved definition of lexical flexibility.

2.2 Traditional approaches

This section is a necessarily brief history of approaches to lexical categories up until the cognitive turn of the 1980s. It covers the universalist position that developed in the Classical tradition, the relativist position that developed as a result of Boas’ cultural relativism, and the structuralist (or “distributionalist”) position that developed in the tradition of Saussure. Depending on how one understands and applies these different perspectives, none of them are mutually exclusive. It is especially common for linguists to simultaneously hold that lexical categories must be identified on the basis of language-internal evidence alone (the relativist position) and that lexical categories are universal in some sense or another (the universalist position).

2.2.1 Universalism

Historically and still presently, many researchers assumed that a small set of lexical categories are basic and universal to all languages (BolingerSears1981Croft 1991: 2Payne1997; Stassen 2011: 95). The set typically consists of some variation of the following: Noun, Verb, Adjective, Adverb, Pronoun, Adposition, Conjunction, Numeral, and Interjection (Haspelmath2001). This set has its origins in the *Τέχνη Γραμματική* / *Tékhnē Grammatiké* (‘The art of grammar’) of the 2nd century B.C.E. grammarian Dionysius Thrax. The *Tékhnē* synthesizes the work of Dionysius’ predecessors, describing eight parts of speech for Ancient Greek. These parts of speech were based largely on morphological (especially inflectional) criteria (Rauh2010). The *Tékhnē* was then translated and its model applied to Latin in the *Ars Grammatica* of Remnius Palaemon. The *Ars Grammatica* initiated a tradition wherein the languages of Europe and eventually the world (e.g. Mandarin [McDonald 2013]) were described using both Dionysius’ categories (with occasionally additions / subtractions) as well as his method of identifying those categories on the basis of morphological criteria (Rauh2010).

Implicit in the Classical method is the assumption that lexical categories are universal in

the sense of being instantiated in all languages. However, as European scholars began to encounter non-Indo-European languages (or even non-Romance languages) in both Europe and abroad, this assumption was challenged, as early as the first grammatical descriptions of Irish in the 7th century. At first, these languages either had Classical grammar imposed upon them or were deemed grammatically deficient (Suarez1983). Nonetheless, missionary linguists in the early colonial era were indeed aware of the significant grammatical differences between these languages and Latin, and made their best attempts at describing them (Suarez1983). It is also important to realize that the project of describing the languages in the Americas and other zones of colonial influence was partially contemporaneous with the publication of the first grammars of the vernacular languages of Europe, as illustrated in Figure 2.1 (the data for which are shown in Table 2.1). Between 1524 and 1572, over 100 catechisms, manuals for confession, collections of sermons, grammars, and vocabularies were written in or about ten languages within the Viceroyalty of New Spain alone (an area smaller than present-day Mexico), mostly by Spanish Franciscan and Jesuit missionaries (Suarez1983). The task of converting the indigenous peoples to Christianity via the medium of their own languages was so important to the Spanish crown that the first bishop of Mexico, Francisco de Zumárraga, brought a printing press to Mexico in 1534 (just 15 years after the arrival of the first Spaniards in Mexico in 1519). The first book printed in Mexico was a Spanish-Nahuatl catechism by Alonso de Molina (Suarez1983). All this is merely to illustrate that language scholars in the colonial era were wrestling with the lexical categories of non-Indo-European languages—and therefore aware of the challenges these languages posed to Classical theories—at a very early stage.

Table 2.1: Some first grammatical descriptions of European vs. American languages

Language	Year	Title	Author
Irish	600s	<i>Auraicept na n-Éces</i> 'The scholars' primer'	Longarad

Table 2.1: Some first grammatical descriptions of European vs. American languages

Language	Year	Title	Author
Occitan	1327	<i>Leys d'amors</i> 'Laws of love'	Guilhèm Molinièr
Welsh	1382–1410	<i>Llyfr Coch Hergest</i> 'Red book of Hergest'	unknown
Tuscan	1437–1441	<i>Grammatica della lingua toscana</i> 'Grammar of the Tuscan language'	Leon Battista Alberti
Castilian	1492	<i>Gramática de la lengua castellana</i> 'Grammar of the Castilian language'	Antonio de Nebrija
French	1530	<i>L'Éclaircissement de la langue fran- coyse</i> 'Explication of the French language'	John Palsgrave
German	1534	<i>Ein Teutsche Grammatica</i> 'A German grammar'	Valentin Ickelsamer
Basque	1545	<i>Linguae Vasconum Primitiae</i> 'First fruits of the Basque language'	Bernard Etxepare
Totonac	1539–1554	<i>Arte de la lengua totonaca</i> 'Grammar of the Totonac language'	Andrés de Olmos
Nahuatl	1547	<i>Arte para aprender la lengua mexicana</i> 'Grammar for learning the Mexican language'	Andrés de Olmos
Tarascan	1558	<i>Arte de la lengua tarasca de Michoacán</i> 'Grammar of the Tarascan language of Michoacán'	Maturino Gilberti
Dutch	1559	<i>Den schat der Duytsscher Talen</i> 'The treasure of the Dutch language'	John III van de Werve
Quechua	1560	<i>Grammatica o arte de la lengua gen- eral de los Indios de los Reynos del Peru</i> 'Grammar or Art of the General Lan- guage of the Indians of the Royalty of Peru'	Domingo de Santo Tomás
Tzeltal Maya	1571	<i>Ars Tzeldaica</i> 'Tzeltal Grammar'	Fray Domingo de Hara
Zapotec	1578	<i>Arte en lengua Zapoteca</i> 'Grammar in the Zapotec language'	Juan de Córdova
English	1586	<i>Pamphlet for Grammar</i>	William Bullokar

Table 2.1: Some first grammatical descriptions of European vs. American languages

Language	Year	Title	Author
Mixtec	1593	<i>Arte de lengua Mixteca</i> 'Grammar of the Mixtec language'	Antonio de los Reyes
Timucua	1614	<i>Gramatica de la lengua Timuquana de Florida</i> 'Grammar of the Timucua language of Florida'	Francisco Pareja
Narragansett	1643	<i>A key into the language of America</i>	Roger Williams

As documentary linguistics turned its attention to North American (as opposed to Mesoamerican) languages, lexical flexibility in particular became a more prominent issue. In fact, even the first comprehensive survey of North American languages contains an entire section on “Conversion of nouns into verbs” (Gallatin1836), in which Gallatin depicts lexical flexibility as a rampant feature of all languages on the continent:

It is the substantive [i.e. copula / auxiliary] verb which we [speakers of Indo-European languages] conjugate; whilst the [Native American] conjugates what we call the adjective and even the noun itself, in the same manner as [s/he] does other intransitive verbs. [...] I believe it must appear sufficiently obvious, that this general if not universal character of the [Native American] languages, the conversion into verbs and the conjugation, through all the persons, tense, and moods, of almost all the adjectives and of every noun which, without a palpable absurdity, is susceptible of it, is entirely due to the absence of the substantive verb. (Gallatin1836)

As evidenced by the above passage, increasing familiarity with non-Indo-European languages prompted some writers to abandon the universalist commitment. However, categorical universalism is still a widely-held position today, either in the sense of a) being universally instantiated in all languages (commonly assumed by most generative frameworks; although see Culicover1999), or b) being available to all languages, but only instantiated in some (sometimes called the “grab bag” approach, as exemplified by Dixon’s Basic Linguistic Theory framework [Dixon2010a; b; c]; [Hieber 2013: ???Croft]).

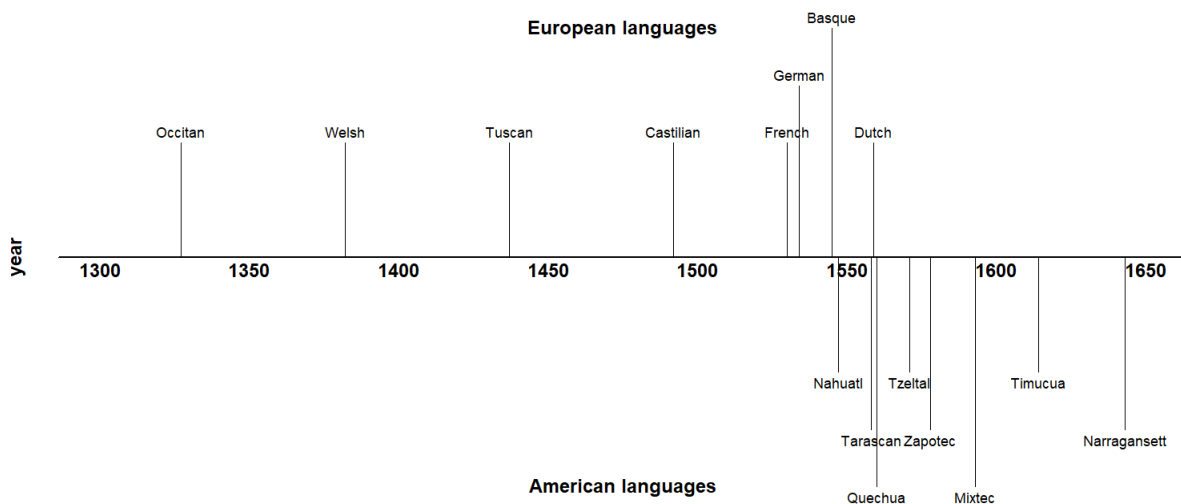


Figure 2.1: Approximate date of some of the first grammatical descriptions of European vs. American languages

2.2.2 Relativism

American ethnographers in the tradition of Franz Boas questioned the universalist assumption in a programmatic and comprehensive way. Writing on grammatical categories, **Boas1911** states, “Grammarians who have studied the languages of Europe and western Asia have developed a system of categories which we are inclined to look for in every language” (**Boas1911**). He concludes that this endeavor is a folly, and that “in a discussion of the characteristics of various languages **different fundamental categories** will be found” (**Boas1911**). Boas’ students all adopted his grammatical relativism, and it became a foundational principle of the American linguistics tradition. His student Edward Sapir, writing on lexical categories specifically, makes one of the best-known and strongest statements of this position in his influential textbook *Language*: “[N]o logical scheme of the parts of speech—their number, nature, and necessary confines—is of the slightest interest to the linguist. Each language has its own scheme. Everything depends on the formal demarcations which it recognizes.” (**Sapir1921**).

Many linguists today hold to Boas’ grammatical relativism in some fashion or another. Textbooks and typological surveys commonly state that languages have varying numbers of

lexical categories, though usually with the caveat that all languages seem to differentiate at least noun and noun (e.g. [Velupillai 2012](#): §6.2). Some researchers, especially those working in typology, argue that linguists are still not rigorous *enough* in their application of grammatical relativism; they criticize certain kinds of crosslinguistic comparisons for imposing the categories of one language onto another ([Croft 2001](#)[Haspelmath2010](#)[LaPolla2016](#)) (CITE: You could probably cite your 2013 article for this too.). This position is discussed further in §2.4.

2.2.3 Structuralism

Developing alongside the early anthropological linguistics of Boas was the linguistic structuralism of Ferdinand de Saussure. His work informed both the Prague school under Nikolay Trubetzkoy and Roman Jakobson, and the distributional method of Leonard Bloomfield. The term *structuralism* has any number of uses ([Matthews2001](#)); here I refer to the idea that “language is a [...] self-contained, self-regulating system, whose elements are defined by their relationship to other elements” ([Matthews2014](#)). In particular, I am referring to the positivistic flavor of structuralism as practiced by Bloomfield, which focused on the structural relations between elements and establishing a set of rigorous scientific discovery procedures for linguistic structures ([Bloomfield1933](#)). Bloomfield saw lexical categories as something to be empirically discovered in the different syntactic distributions of words, rather than imposed on a language a priori ([Rauh2010](#)). Zellig Harris later refined and expanded on this methodology ([Harris1951](#)), which in turn was incorporated into Noam Chomsky’s ealy Phrase Structure Grammar.

The signature methodological feature of this form of structuralism is the *distributional method*, a procedure for defining categories in terms of the set of contexts in which its words can appear—that is, their distributions ([Harris1951](#)[2001](#): 11). As an illustration of distributional analysis applied to lexical categories, [Croft \(1991](#): 11–12) considers the distributions of the English words *cold*, *happy*, *dance*, and *sing* in two constructions: in the Predicate construction after *be*, and in the 3rd Person Singular Present Tense (-s) construction. Example

data are shown below.

(11) English (Indo-European > Germanic)

- a. i. Jack is cold.
ii. * Jack colds.
- b. i. Jack is happy.
ii. * Jack happies.
- c. i. * Jack is dance.
ii. Jack dances.
- d. i. * Jack is sing.
ii. Jack sings.

We can see that *cold* and *happy* have the same distributions in these tests (both may appear in the Predicate construction but not the Person-Tense inflection construction), while *dance* and *sing* have the same distribution (the inverse situation as *cold* and *happy*). The results of these two distributional tests are summarized in [Table 2.2](#).

Table 2.2: Distribution of English Verbs and Adjectives (adapted from [Croft \[2001: 12\]](#))

	Predicate Construction	Inflectional Construction
Adjective: <i>cold, happy</i> , etc.	✓	✗
Verb: <i>sing, dance</i> , etc.	✗	✓

As applied in practice, however, the distributional method suffers from one serious drawback when used to argue for large, traditional categories like noun, verb, and adjective: distributional tests yield conflicting and overlapping results. Perhaps no two lexical items behave the same in every distributional test. Each new test that is introduced therefore partitions the lexicon into smaller and smaller classes. This fact has been demonstrated empirically for English temporal nouns ([Crystal 1967: 54](#)), Russian numerals ([Corbett1978](#)), and French verbs ([Gross1979](#)). Distributional tables like that in [Table 2.2](#) from each of these studies are reproduced in Table ??, Table ??, and Table ?? respectively. It is clear from these studies that distributional analysis does *not* lead to large, unified categories like noun, verb, and adject-

tive, but rather a myriad of small constructions (Crystal 1967: 27; Croft 2005: 434). Each distributional test is in fact its own construction (Croft 2005: 436). This fact is the motivation underlying constructional approaches to language.

Many scholars nonetheless choose to retain lexical categories as a necessary component of their linguistic theories or descriptions, at the expense of consistent application of the distributional method. Rather than considering all possible distributional contexts for a word, these scholars instead treat certain constructions as definitional of the category. Other distributional tests which yield cross-cutting results are either ignored, or treated as evidence of subcategories instead of categories. Many researchers even prefer the term *syntactic categories* over *lexical categories* for this reason, focusing on just the syntactic evidence for categories (CITE: Baker 2003; Rauh 2010). A severe methodological problem for this approach is that there are no generally agreed-upon principles for determining which distributional tests should be considered definitional. In this regard, Schachter & Shopen (2007: 4) note, “there may be considerable arbitrariness in the identification of distinct parts of speech rather than subclasses” (see also Crystal [1967]). Different scholars choose or prioritize different kinds of evidence for lexical categories over others on the basis of their theoretical commitments. This is the reason, as stated in Section 1.1, that disagreements about the existence of particular lexical categories in particular languages are typically *not* about the empirical facts. The results of a given distributional analysis are not usually controversial; the choice of distributional tests used to support one’s analysis are. Unsurprisingly, then, debates over how to analyze lexical categories in various languages have been largely unproductive and unresolved (Croft 2005: 435). The problem only worsens when scholars attempt to apply the same criteria across languages. Distributions of words with similar meanings vary drastically across languages (Croft 2001: §1.4.1).

The real methodological problem here is *not* that we have yet to ascertain the correct principles for selecting the right distributional tests. The problem is being selective regarding which tests to apply in the first place. If we take the distributional method seriously,

	in a <i>N</i> or two	in that <i>N</i>	in the <i>N</i> _{sg.} (no postmodification) ⁶⁵⁾	in a <i>N</i> (no postmodification) ⁶⁵⁾	in \emptyset <i>N</i> _{pl.} (no postmodification) ⁶⁵⁾	in \emptyset <i>N</i> _{sg.}	\emptyset <i>N</i> _{pl.} on the <i>N</i> _{sg.}	on a <i>N</i> _{sg.} (no postmodification) ⁶⁵⁾	on \emptyset <i>N</i> _{pl.} (no postmodification) ⁶⁵⁾	on \emptyset <i>N</i> _{sg.}	at that <i>N</i>	at the <i>N</i> _{sg.}	at \emptyset <i>N</i> _{sg.}
afternoon	+	+	+	+	-	-	+	+	+	-	-	-	-
evening	+	+	+	+	-	-	+	+	?+	-	-	-	-
weekend	+	+	?+	+	-	-	+	+	?+	-	+	+	-
night	+	+	+	?+	-	-	?-	+	-	- ⁶⁶⁾	-	-	+
morning	+	+	+	+	-	-	+	+	-	- ⁶⁶⁾	-	-	-
Monday...	+	-	-	-	-	-	+	+	+	+	-	-	-
January...	?+	+	+	+	-	+	-	-	-	-	-	-	-
hour	+	+	+	+	?-	-	-	+	-	-	+	+	-
minute	+	+	-	+	+	-	-	+	-	-	+	+	-
second	+	+	-	+	+	-	-	+	-	-	+	+	-
day	+	+	+	+	-	-	-	+	-	- ⁶⁶⁾	-	-	-
summer	?+	+	+	?-	?--	+	?+	-	-	-	-	-	?+
winter	?+	+	+	?-	?--	+	?+	-	-	-	-	-	?+
spring	?+	+	+	?-	-	+	-	-	-	-	-	-	?+
autumn	?+	+	+	?-	-	+	-	-	-	-	-	-	?+
month	+	+	+	+	+	-	-	-	-	-	-	-	-
week	+	+	+	+	+	-	-	-	-	-	-	-	-
year	+	+	+	+	+	-	-	-	-	-	-	-	-
decade	+	+	-	+	?+	-	-	-	-	-	-	-	-
century	+	+	-	+	?-	-	-	-	-	-	-	-	-
fortnight	+	+	+	+	-	-	-	-	-	-	-	-	-
instant	+	?+	-	+	-	-	-	-	-	-	+	-	-
moment	+	?+	-	+	-	-	-	-	-	-	+	-	-
lifetime	-	?+	-	+	-	-	-	-	-	-	-	-	-
daytime	-	-	+	-	-	+	-	-	-	-	-	-	-
nighttime	-	-	+	-	-	+	-	-	-	-	-	-	-

Figure 2.2: distributional analysis of English temporal nouns (Crystal 1967:

54)

	<i>odin</i> 1	<i>dvo</i> 2	<i>tri</i> 3	<i>pjat'</i> 5	<i>sto</i> 100	<i>tysjača</i> 1,000	<i>million</i> 1,000,000
1. Agrees with N in syntactic number	+	–	–	–	–	–	–
2. Agrees in case throughout	+	–	–	–	–	–	–
3. Agrees in gender	+	(+)	–	–	–	–	–
4. Marks animacy	+	+	+	–	–	–	–
5. Has own plural	+	+	+	+	(–)	–	–
6. Takes agreeing determiner	+	+	+	+	+	–	–
7. Takes N in genitive plural throughout	+	+	+	+	+	±	–

Figure 2.3: distributional analysis of Russian numerals (Corbett1978)

then we must apply it consistently, without ignoring distributional evidence that contradicts our theoretical or pretheoretical assumptions. To do otherwise is a kind of *methodological opportunism* (CITE: Croft 1991: 10–15, 41 (check these pages); Croft 2001: ???).

add a section on mixed / hybrid categories (including Lois, Malouf, Nikolaeva)

Partly in response to these problems, a growing cadre of linguists in the last 30 years have adopted one of various *flexible* approaches to word classes. Flexible analyses of word classes come in many flavors, some of which arguably still commit methodological opportunism, and others of which introduce new difficulties. These flexible approaches are reviewed in the following section.

2.3 Flexible approaches

In this section I summarize the key concepts (§2.3.1), findings (§2.3.2), and criticisms (§2.3.2) of lexical flexibility research. Section 2.3.1 surveys the wide variety of definitions and theoretical perspectives on lexical flexibility. This review of the literature reveals that there is very little consensus as to what exactly constitutes “lexical flexibility”; as such, there are nu-

il V Ω					Compléments directs ou indirects										Comp. indirect												
					Auxiliaire avoir N_0 est Vpp Ω N_0 V Prép N_1					Complétives					Noms							Noms					
										que P	que Psubj [pc z.]	V ² Ω	ce (ci + la)	ppv	N_{hum}	ppv	N_{hum}	le fait Qu P	ppv			N_0 V Prép N_2	N_{hum}	ppv	N_{hum}	ppv	N_0 V V ⁰ Ω
+	s'agir	-	+	-	+	de	-	+	+	+	+	-	+	-	+	+	-	pour	+	-	-	-					
+	apparaître	-	+	-	+	0	+	-	-	-	+	+	-	-	+	-	+	-	à	+	+	-	-				
+	apparoir	-	-	-	-	0	-	+	-	-	-	-	-	-	-	-	-	0	-	-	-	-					
+	s'avérer	-	+	-	+	0	+	-	-	-	+	+	-	-	+	-	+	-	à	+	-	+	+				
+	y avoir avantage	+	-	-	+	à	-	+	-	+	+	-	-	-	-	-	+	pour	+	-	-	-					
+	y avoir lieu	+	-	-	+	de	-	+	+	+	-	-	-	-	-	-	-	pour	+	-	-	-					
+	n'empêcher	+	-	-	+	0	+	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-					
+	être besoin	+	-	-	-	de	-	+	+	+	+	+	-	+	-	+	-	pour	+	-	+	-					
+	être l'heure	+	-	-	-	de	-	+	+	+	+	+	-	+	-	+	-	pour	+	-	-	-					
+	être question	+	-	-	+	de	-	+	+	+	+	+	+	-	+	+	+	pour	+	-	+	-					
+	être temps	+	-	-	+	de	-	+	+	+	+	+	-	+	-	+	-	pour	+	+	-	-					
+	faire bon	+	-	-	+	0	-	-	-	+	-	-	-	-	-	-	-	pour	+	-	-	-					
+	falloir	+	-	-	+	0	-	+	-	+	+	+	+	+	-	+	-	à	+	+	-	-					
+	s'en falloir	-	+	-	-	pour	-	+	-	-	-	-	-	-	-	-	-	de	-	-	+	-					
+	paraître	+	-	-	+	0	+	-	-	-	+	+	-	+	-	+	-	à	+	+	-	+					
+	paraître	-	-	-	+	0	+	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-					
+	sembler	+	-	-	+	0	+	-	-	+	+	-	-	+	-	+	-	à	+	+	-	+					
+	souvenir	-	+	-	+	de	+	-	+	+	+	+	-	+	+	+	-	à	+	+	-	-					

Figure 2.4: distributional analysis of French verbs (Gross1979)

merous alternative terms for the phenomenon. Despite these incongruities, a few important findings do consistently surface across the empirical research. These findings are summarized in [Section 2.3.2](#). [Section 2.3.3](#) then looks at the arguments and evidence that researchers have presented against the notion of lexical flexibility.

2.3.1 Key concepts

It is only a small exaggeration to say that there are as many definitions and terms for what I am here calling “lexical flexibility” as there are scholars who research it. I use the term *lexical flexibility* in this thesis merely because it is the most widely recognized of the cluster of terms that are used, not because it is necessarily the most precise or accurate. My own choice would be *functional expansion*, for reasons discussed below. The analytical or theoretical perspective adopted by each researcher generally determines their choice of terminology. The remainder of this section is devoted to explaining these perspectives in more detail.

Generally speaking, there are two ways to analyze flexible words. The first method assigns flexible words to members of specific categories in a language, whether those language-categories are the canonical four major classes (Noun, Verb, Adjective, Adverb), or a new large supercategory subsuming multiple discourse functions (e.g. contentives, non-verbs, flexibles), or a smaller subcategory of an existing major lexical category (e.g. adjective verbs, verbonominals). The second method of analysis assumes that words are uncategorized at some level (root, stem, or inflected word), and that words receive their categorial assignment from context. Different researchers posit different mechanisms for how words receive their categorization in context. The traditional approaches to lexical flexibility summarized in [Section 2.2](#) are all instances of the former method of analysis, while the flexible approaches outlined in this section are a mix of categorial and non-categorial analyses.

2.3.1.1 Lexical flexibility

Though awareness of lexical flexibility can be traced back to at least Gallatin 1836 if not earlier, the term *lexical flexibility* itself seems to have originated with Hengeveld 1992. This publication, perhaps because it was the first to assign a technical term to the concept, marks a shift in how scholars frame the concept of lexical flexibility. Previously, the issue was framed in terms of whether particular languages (especially those of the Pacific Northwest) distinguished noun from verb (Kuipers 1968; Jacobsen, Jr. 1979; Hébert 1983; Kinkade 1983; van Eijk & Hess 1986; Jelinek & Demers 1994). After this point, an increasing number of publications began to ask whether lexemes were *flexible* instead. Though the difference in emphasis seems subtle, this change constitutes a turning point because it fostered an increased interest in lexical flexibility as a grammatical phenomenon in its own right instead of just a problem for traditional categorization schemes.

Hengeveld 1992's Hengeveld 1992 typology of parts-of-speech systems is a whole-language typology wherein languages are either *specialized*, with one morphosyntactic category for each of the functions of reference (Noun), predication (Verb), referent modification (Adjective), and predicate modification (Adverb¹), or *non-specialized*. Non-specialized languages deviate from the four-category canon in one of two ways: one part of speech may assume more than one function with no additional morphosyntactic marking, in which case the language is considered *flexible*; or the language may lack a dedicated part of speech for that function entirely and use other, marked constructions instead, in which case the language is considered *rigid*.

add examples of flexible vs. rigid languages, with explanations

Hengeveld 1992's analysis is of the categorial type discussed at the beginning of [Section 2.3.1](#), specifically the supercategory kind. Each word is assumed to have a category, and new supercategories are introduced for words which have multiple functions: *contentives*

¹Note that Hengeveld's typology only includes manner adverbs, not other semantic types of adverbs.

for words which perform all four functions, *non-verbs* for words which perform all non-predicating functions, and *modifiers* for words which perform referent modifier and predicate modifier functions.

Hengeveld1992's parts-of-speech typology and the subsequent research it inspired (DonLier2003Hengeveld & Rijkhoff 2005Lier2006HengeveldLier2010; Luuk 2010LierRijkhoff2013Lier2016) constitute important empirical contributions to the study of lexical flexibility. However, Hengeveld's definition of flexible languages and his parts-of-speech typology still rely on large, language-specific categories of the kind that have been problematized by Croft (CITE: Croft's publications, especially the pages in 1991 where he critiques Hengeveld directly) and van Lier (CroftLier2012), and are therefore subject to the same difficulties as traditional approaches to parts of speech. However, numerous scholars have since adopted Hengeveld's term *lexical flexibility* to describe cases where words serve more than one discourse function, regardless of their particular theoretical commitments or analysis of flexible words. As a convenient cover term, *lexical flexibility* is now well established.

2.3.1.2 Polycategoriality

VapnarskyVeneziano2017 introduce the alternative term *polycategoriality* as their preferred characterization of flexible words. (The term is also used by Carter2006, but he does not give a precise definition for it.) While VapnarskyVeneziano2017 use this term mostly interchangeably with *lexical flexibility*, there are important differences between the two concepts. Hengeveld1992's use of *lexical flexibility* is meant to imply the existence of large, flexible supercategories that subsume multiple discourse functions, whereas Vapnarsky & Veneziano are not committed to any particular schema for parts of speech. Central to their notion of polycategoriality is the idea that lexical categories exist, but that "there are lexical forms that are not specified for lexical category (or are not specified fully, or univocally) on some level of representation." (VapnarskyVeneziano2017). In other words, one word may belong simultaneously to multiple lexical categories. Under this definition, a language could

still have all four major lexical categories but nonetheless exhibit rampant polycategoriality; this is not a possibility in Hengeveld1992's framework. Like Hengeveld1992, however, VapnarksyVeneziano2017 are committed to the existence of large lexical categories in particular languages. Their analysis is therefore also of the categorial kind discussed at the beginning of [Section 2.3.1](#).

2.3.1.3 Multifunctionality / Polyfunctionality

Another term for our phenomenon of interest, introduced by (Lier2012), is *multifunctionality*, in which a single lexeme can have multiple discourse functions. An advantage of this analysis is that it takes no theoretical position on the issue of whether words are categorial or non-categorial; it focuses on just their functions. The term *multifunctionality* is meant to stand in contrast with *conversion* or *zero derivation*. Lier2012 takes conversion to be idiosyncratic and unproductive, producing meanings for words in alternate discourse functions that are not predictable (see [section on conversion](#) and [section on item-specific knowledge](#) for further discussion). Multifunctionality is also distinct from zero derivation from a common root. Instead, multifunctional words are those whose semantic interpretation is entirely predictable from context, and its uses in different contexts are productive. Its meaning should be *compositional* (see [section: criteria for flexibility > compositionality](#)). For example: when an action word is used in a referring construction its predicted meaning is that of an *action nominalization*, '(the act of) X-ing'; and when an object word is used in a predicate construction its predicted meaning is that of a *predicate nominal*, 'be an X'. Examples of these predictable, compositional meanings are shown in [EX and EX](#).

add examples of multifunctionality

However, words used in their non-prototypical functions very frequently do not have predictable meanings. Consider the examples in [EX and EX](#).

add examples of conversion

Add lead-out explaining how the meanings of the latter examples are idiosyncratic and unpredictable.

Lier2012 takes examples like those in *former examples* to be instances of genuine multifunctionality, and those in *latter examples* to be cases of conversion. Others have also adopted a position similar to van Lier's, in which only the semantically compositional / predictable uses of a word in different discourse functions are considered flexible (Evans & Osada 2005: pages for compositionality criterion) (CITE: Croft 2005? Where he is critical of flexible analyses for ignoring idiosyncratic meanings).

2.3.1.4 Precategoriality / Acategoriality

The various approaches which analyze flexible words as being at some level uncategorized until they receive their interpretation from context may be lumped together under the umbrella terms *precategoriality* or *acategoriality*. Hopper & Thompson's influential paper 1984 is an early application of the concept of acategoriality to the analysis of flexible words:

[L]inguistic forms are in principle to be considered as *lacking categoriality* completely unless nounhood or verbhood is forced on them by their discourse functions. To the extent that forms can be said to have an apriori existence outside of discourse, they are characterizable as *acategorial*; i.e. their categorial classification is irrelevant. Categoriality-the realization of a form as either a N or a V-is imposed on the form by discourse. (Hopper & Thompson 1984: 747)

The term *precategorial* has become a somewhat more common term for roughly the same concept (though some researchers use the term in more strictly-delineated ways) (Evans & Osada 2005: 357, 362–364; Bisang2008a; Bisang2008b; Bisang2013). It is especially preferred by morphological models that presuppose stages of derivation, such that words are precategorial before they reach a certain stage of the derivation (Halle & Marantz1993; Arad2005) (CITE: chapter from Handbook of morphological theory). Vapnarsky & Veneziano2017 distinguish polycategoriality from acategoriality by defining acategoriality as implying “no primitive / original categorial marking at all”, and polycategoriality as allowing a form “to be only partially unspecified for category, with possible constraints on the relevant categories”. Languages for

which precategorical analyses have been advanced include Cherokee (Haag2017), Gooniyandi (McGregor2013), Kuikuro (FranchettoSantos2017), Mundari (Hengeveld & Rijkhoff 2005), and many others. Pfeiler2006 also presents psycholinguistic evidence that the earliest utterances of L1 learners of Yucatec Maya are acategorical.

A central concern in precategorical approaches is the precise mechanism by which a word receives its categorization in context (HengeveldRijkhoffSiewierska2004). There are two main theories of semantic indeterminacy in flexible words: *underspecificity* (Farrell 2001RijkhoffLier2013) and *vagueness* (Tuggy1993HengeveldRijkhoffSiewierska2004Hengeveld & Rijkhoff 2005). The essential difference is that underspecificity entails semantic minimalism while vagueness is entails semantic maximalism. An underspecified word has a minimal, core meaning, and receives its categorial meaning from the discourse context it appears in; a vague word has a maximal, broad meaning that covers all the possible discourse contexts it appears in. (There is of course quite a deal of variation in the literature as to how scholars use these terms, with many researchers conflating the two.) Hengeveld & Rijkhoff (2005: 414) offer the example of *cousin* as a word that is semantically underspecified for gender, such that the gender of the referent must be understood from context. Denison2018 argues that the English word *long* exhibits adjective ~ adverb underspecification in Old and Middle English.

In contrast, HengeveldRijkhoffSiewierska2004 outline a theory regarding exactly how vagueness operates in the context of lexical flexibility:

[E]ach flexible lexeme has a single (vague) sense. By placing the flexible lexeme in a particular syntactic slot or by providing it with certain morphological markers, the speaker highlights those meaning components of the flexible lexeme that are relevant for a certain lexical (verbal, nominal, etc.) function. Thus we contend that the meaning of a flexible lexeme always remains the same, and that morphosyntactic and other contextual clues signal to the addressee how to interpret this lexeme in an actual utterance. In other words, it is the use of a vague lexeme in a certain context (an actual linguistic expression) that brings out certain parts of its meaning, giving the category-neutral lexeme a particular categorial (verbal, nominal, etc.) flavour. (Hengeveld & Rijkhoff 2005: 541)

(Note that while vagueness implies a certain potential ambiguity, Hengeveld & Rijkhoff reserve the term *ambiguity* for cases of distinct, homophonous words.)

Evans & Osada (2005: 363–364) and Kihm2017 criticize both precategorical approaches for their imprecision, claiming that it would be impossible to formulate a definition for many flexible words that was broad enough to encompass all their uses. Kihm2017 illustrates this difficulty with the various meanings of the Arabic root *s-q-t*, which could arguably be glossed FALL. A selection of stems containing this root are given in (11).

add examples from Kihm (2017: 87)

It is difficult to imagine a single definition of *s-q-t* which could adequately demarcate just this set of meanings. This difficulty could perhaps be overcome, however, by loosening the requirement that the meaning of a word be unitary. Word meanings are *polycentric*, where different senses of a word are related through *meaning chains* rather than all through a single, central member (Taylor 2003: 110). This is sometimes referred to as a *family resemblance* structure for categories. The difference between monocentric and polycentric categories is illustrated schematically in Figure ???. In both diagrams, each letter A–E represents a sense of a word. In the monocentric case, all the senses of the word are related through its core sense A. In the polycentric case, senses A and E are related only through their intervening connections.

Create a 2-part figure comparing monocentric vs. polycentric categories, with monocentric on the left. The senses in the monocentric case should be arranged circularly around sense A. The senses in the polycentric case should be arranged linearly from A > E.

Recognizing that word meanings are polycentric addresses Evans & Osada and Kihm2017's criticisms of vagueness theory because it shows that the disparate senses of a word can be related without having to share any core component of their meanings. The use of a lexeme in a certain context then profiles one of these senses over others. Kihm himself hints at this solution by referring to the related Arabic stems in (11) as a *lexical family*.

2.3.1.5 Monocategoriality

In the extreme case where all lexical words in a language are precatgorial, the language could be considered *monocategorial*, possessing a single, open syntactic category. This is effectively the same as saying that the language lacks lexical categories altogether, the difference being primarily one of emphasis. David Gil analyzes both Tagalog (1993Gil1995) and Riau Indonesian (1994) as being of this extreme monocategorial type. Moreover, he argues that monocategoriality must have been typical of an earlier stage of language evolution in which dedicated morphosyntactic constructions for different discourse functions had yet to evolve (Gil2005aGil2005Gil2012). He names this abstract language type an *Isolating-Monocategorial-Associational* (IMA) language.

2.3.1.6 Transcategoriality

It is also worth briefly mentioning *transcategoriality*, since the term arises occasionally in connection with lexical flexibility and is potentially easily confused with other terms mentioned above. Robert2003 uses *transcategoriality* to describe the ability for a single form to serve both lexical and grammatical functions. This is common in grammaticalization scenarios in which the original, lexical use of a form continues to exist alongside its newer, functional use. This is commonly referred to in the grammaticalization literature as *divergence* (Hopper & Traugott 2003: 118). Since the focus of lexical flexibility is on *lexical* words and categories rather than *functional* ones, the concept of transcategoriality is not directly relevant to the study of lexical flexibility.

2.3.1.7 Conversion / Zero derivation

Since the literature on conversion and zero derivation is extensive and the concepts are well-established, I will treat them only summarily here, focusing on their relationship to lexical flexibility. *Conversion* is the process whereby a lexical item simply changes its word class with no overt morphological marker of that change (Crystal2008). *Zero derivation* is an alternate

analysis of the same phenomenon that posits the presence of a derivational marker with no phonological realization. (CITE: Add citations to either conversion or zero derivation as appropriate: Bauer & Varela 2005; Don 2005; Kerleroux 1996, 2000; Lieber 2006; look at other handbooks of typology, morphology, and word classes) I will refer to this process of unmarked shift from one category to another as *conversion* in this thesis.

The concept of conversion is based on the premise that words are fully categorized for part of speech, meaning that an analysis of lexical flexibility as conversion falls under the categorical (as opposed to non-categorical) analyses of lexical flexibility mentioned at the beginning of [Section 2.3.1](#). Conversion is generally characterized as a kind of word formation, implying that a new word has been created. Therefore, conversion and lexical flexibility are mutually exclusive analyses of multifunctional words; lexical flexibility implies the existence of one polysemous word which can fulfill multiple discourse functions, while conversion implies the existence of two homonymous / heterosemous words with different discourse functions. Remember too from [Section 2.3.1.3](#) that [Lieber 2012](#) distinguishes conversion from multifunctionality, where conversion is reserved for unproductive / unpredictable derivations. Not all scholars would delimit conversion in this way however.

Conversion also implies directionality. In cases of conversion, one of the two uses of a form is in some way basic or prior to the other ([Mithun 2017](#): 156 [Vapnarsky & Veneziano 2017](#)). Under a flexible analysis, by contrast, the different functions of a single flexible lexeme have equal theoretical status. If it could be shown that certain seemingly flexible uses of a word were in some way marked in relation to each other, this would therefore constitute potentially disconfirming evidence against a flexible analysis. This is in fact one of the major arguments presented against flexible analyses, to be discussed in [Section 2.3.3](#). There are at least four ways in which one member of a putatively flexible set of lexical items might be considered more basic than the others: diachronically, in which one sense of the word appears before the others historically; semantically, in which the meaning of the derived word is more semantically complex than that of the basic word; morphologically, in which the more basic

word is irregularly inflected but the derived word is regularly inflected; or frequently, in which derived words are used less frequently than their base words (Plag2003). Speakers themselves also have intuitions about which member of a flexible set is basic and which are derived (Mithun 2017: 166). As will be explained in Section ??, the idea that certain uses of a form are marked in relation to each other is also central to Croft's typological markedness theory of lexical categories.

2.3.1.8 Functional shift / Functional expansion

Especially in the American context, another common term for conversion is *functional shift* (Cannon 1985). In most research, the term is used essentially interchangeably with *conversion* or *zero derivation*. However, functional shift can be usefully distinguished from conversion by its emphasis on function over than category, paralleling the distinction between polycategoriality (implying language-specific categories) and polyfunctionality (with no such implication). In its literal interpretation, the term suggests a shift in the meaning of a word from one discourse function to another, an analysis amenable to a constructional approach, and one that is not committed to the existence of language-particular categories. A slight improvement on this term would be *functional expansion*, since it emphasizes the expansion of a linguistic form into new functions / contexts as opposed to the wholesale shift from one function to another implied by *functional shift*.

2.3.2 Key findings

The emergence of lexical flexibility as an object of study has led to a number of edited collections or journal volumes focused on flexibility and word classes more generally (VogelComrie2000, Evans & Osada 2005 (target article), Ansaldopfau2010, LoisVapnarsky2010, RijkhoffLier2013, Simone2014, BlaszcakKlimekJankowskaMigdalski2015, VapnarsyVeneziano2017, Lier2017 (target article), CuyckensHeyvaertHartmann2019), plus any number of individual articles (see especially Farrell [2001], Rijkhoff [2007], Lier2012, and Mithun2018). Out of these col-

lections have emerged several recurring findings, each of which is summarized in this section.

2.3.2.1 Parts-of-speech hierarchy

In addition to laying out a theory of flexible categories, **Hengeveld1992** also presents the results of a 30-language survey of parts of speech in which he finds that the categories which are most likely to occur as an independent class in a language are subject to an implicational hierarchy, shown in (12), which Hengeveld refers to as the *parts-of-speech hierarchy*.

(12) Verb > Noun > Adjective > Adverb

Categories to the left of the hierarchy are more like to occur as a distinct part of speech than categories to the right. Applying this hierarchy to **Hengeveld1992**'s flexible vs. rigid distinction yields the parts-of-speech typology in Figure ?? (adapted from **Hengeveld1992** and **Rijkhoff [2007: 718]**). The terms for the different categories in flexible languages are from **HengeveldRijkhoffSiewierska2004**. **Hengeveld1992** points out that this is not a strict classification scheme; languages may sit at the boundaries between types and exhibit exceptions.

parts of speech typology from Hengeveld & Rijkhoff 2005: 407

As mentioned in [Section 2.3.1.1](#), **Hengeveld1992**'s typology could be criticized for its reliance on large, language-specific lexical categories instead of constructions. One could however reframe **Hengeveld1992**'s implicational hierarchy in terms of functions rather than categories, as in (13). I call this the *hierarchy of discourse functions*.

(13) predicate > referent > predicate modifier > referent modifier

In (13), functions to the left of the hierarchy are more likely to have dedicated morphosyntactic constructions than those to the right. This reformulation avoids a commitment to any language-particular categories while still capturing the important implicational trend observed by Hengeveld.

This hierarchy of discourse functions has proven to be a fairly robust finding in the literature on lexical flexibility, now supported by a number of subsequent studies (CITE: all the

various sources that apply and extend the hierarchy of discourse functions).

2.3.2.2 Reference-predication asymmetries

The hierarchy of discourse functions also hints at another important feature of lexical categories: there is something privileged about the predicating function. A survey of the literature on lexical flexibility reveals patterned asymmetries in the behavior of lexical items with regard to predication vs. reference, even in very flexible cases. For starters, while it is quite common for languages to freely allow object words to be used as predicates with no special marking, the reverse case is much less likely [Hopper & Thompson \(1984: 745\)](#). The functional expansion of a word's uses from predication into reference always seems to be more marked (or at least as marked) as the shift from a referring function to a predicating function.

This fact has been observed independently by numerous researchers. For example, [Stevick 1968](#) and [Marchand 1969](#) both observe that conversion from noun to verb in English has always been more common than from verb to noun, and [Kastovsky \(1996: 98\)](#) points out that English does not even have a native noun > verb derivational suffix—any affixes of this type are borrowed from Romance languages. Central Alaskan Yup'ik is another example of a language with very many nominalizers but few verbalizers ([Mithun 2017: 158](#)).

Flexibility itself is frequently *unidirectional*, meaning that any object word may be used for predication, but that action words used for reference are marked ([Croft 2001: 69](#)) [\[CITE:](#)

[Beck's works on unidirectional flexibility; Evans & Osada on bidirectionality\]](#). [Nakayama 2003](#) frames flexibility in Nuuchahnulth in terms of a word's ability to predicate, reporting that “all inflectional stems are potentially predicative”, but the reverse is not true. Discussing Classical Nahuatl, [Launey \(1994; 2004\)](#) introduces the term *omnipredicativity* to describe languages in which all words are potentially predicative, but no corresponding term *omnireferentiality* has appeared in the literature. That said, languages which have undergone *insubordination* (in which subordinate, usually nominalized clauses are reanalyzed as main clauses, and the nominal inflectional marking of the subordinate clause is reinterpreted as verbal inflectional

marking) (CITE: Mithun's 2008 article; Evans' original article; introduction to Evans' edited volume) do exhibit noun-oriented flexibility in the sense that verbal inflection mirrors nominal inflection. This famously led to the claim that all words in Eskimo languages are fundamentally mentally nominal in nature (CITE: Sadock's nominalist theory of Eskimo). However, cases of insubordination do not constitute counterexamples to the predicating tendency in language. Even in these languages, the use of action words for reference is still less marked than the use of object words for predication; this includes Eskimo, as mentioned for Central Alaskan Yup'ik above.

Kastovsky (1996) argues that this asymmetry arises from the fact that “deverbal nouns have a much more diversified semantics than denominal verbs” (Kastovsky 1996: 96), meaning that the range of possible meanings for a deverbal noun (a noun derived from a verb) is significantly broader than for a denominal verb (a verb derived from a noun). Examining data from English, Kastovsky shows that when an object word is used to predicate, its possible meanings are limited to combinations of BE, BE LIKE, BE IN, BECOME, HAVE, DO, DO WITH, and CAUSE. When an action word is used as a predicate, however, the range of meanings include any abstract representation of the event itself (an action nominalization), or any one of the arguments associated with the verb, which come in a variety of semantic roles.

A similar, cognitively-oriented explanation for reference-predication asymmetries is given by Hopper & Thompson (1984: 745):

[A] nominalization names an event taken as an entity; however, a “verbalization” does not name an “entity taken as an event”, but simply names an event associated with some entity. In other words, a nominalization still names an event, albeit one which is being referred to rather than reported on in the discourse; it is, accordingly, still in part a [verb], and not a *bona fide* [noun]. However, a denominal [verb] no longer names an entity at all, and thus has no nominal “stains” to prevent its being a *bona fide* [verb]. (Hopper & Thompson 1984: 745)

Hopper & Thompson (1984: 746) analyze nominalizations as a kind of metaphor following Lakoff & Johnson 1980, in which an abstract event is conceptualized as a concrete object. However, no such metaphor exists for verbalizations, explaining the asymmetry in the direc-

tionality of lexical flexibility.

2.3.2.3 Locus of categoriality

The grammatical level at which a language exhibits flexibility—the root, the stem, or the fully-inflected word—differs from one language to the next. In some languages, roots are strongly associated with a particular discourse function, but stems are flexible; in other languages, the reverse is true. I refer to the linguistic level at which a language associates different discourse functions as its *locus of categoriality*. Some linguistic theories include a premise that the locus of categoriality in every language always sits at a certain level (HalleMarantz1993Baker2003Baker2015BooijAudring2018; Siddiqi 2018), but the evidence from research on lexical flexibility gives strong empirical support to the position that locus of categoriality varies from language to language. In contrast, BlaszcakKlimekJankowskaMigdalski2015 argue that category information is distributed across different levels of representation.

As one illustration of how flexibility depends on grammatical level, we have seen that roots in Central Alaskan Yup'ik are generally categorial—except for 12% of roots, they are typically strongly associated with just one discourse function, and derivational affixes select for roots of a particular category (Mithun 2017: 162–167). While many derived stems are also strictly associated with just one discourse function, a large but indeterminate number have both referential and predicative uses. Some examples of categorial derived stems are in EX and examples of flexible stems are in EX. See also the examples in (7) in Section 1.1.

add examples of from Mithun (2017: 167–168)

Fully-inflected words in Central Alaskan Yup'ik, however, never exhibit flexibility (Mithun2019). So Central Alaskan Yup'ik displays partial flexibility at the root and stem level but not the inflected word level.

As another example, in Mandinka all stems are flexible. No Mandinka stem except for *săa* ‘die’ is used in just one discourse function (Creissels 2017: 46). At the level of the inflected word, however, lexical items in Mandinka belong unambiguously to one category or another

(Creissels 2017: 37). Mandinka therefore shows complete flexibility at the stem level but complete rigidity at the inflected word level. (Creissels does not include an analysis of roots in his discussion.)

Surprisingly, some languages display flexibility even at the level of the fully-inflected word. In many North American languages, it is common for fully morphological verbs to function as referents (Hieber forthcoming), as shown in the following examples.

add examples of flexibility at the word level from p. 20 of your word classes chapter: Cayuga, Chitimacha, Navajo

Each of these flexible uses of a morphological verb sits somewhere on a continuum between being fully lexicalized as a referent, so that its predicating use is no longer available, to being a fully productive predicate, with both predicative and referential uses (Mithun 2000: 413).

The reason that words may exhibit flexibility at one level of analysis but not another is because “*categorical* shift is often not *categorical*” (Mithun 2019). When a word expands its use into new contexts, not all the morphological, syntactic, and semantic properties of the word shift to accommodate that new use at the same time. It takes some time before the morphosyntactic properties of a word adjust to reflect its new use, a process referred to as *actualization* in the grammaticalization literature (DeSmet 2012) and *post-constructionalization* *constructional changes* in the framework of diachronic construction grammar (Hopper & Traugott 2003: 27).

It is because the locus of categoriality can vary from language to language that I have used the vague and overloaded terms *word* and *lexical item* throughout this thesis. Both are intended to be convenient cover terms for root, stem, or inflected word.

2.3.2.4 Item-specificity

A final significant finding to emerge from the empirical research on lexical flexibility is the fact that flexibility is item-specific and even sense-specific. Individual lexical items or even individual senses of an item differ in their flexibility. To say that lexical flexibility is *item-specific* is to say that two words that are otherwise very similar in their meanings and morphosyn-

tactic behavior can nonetheless differ in terms of their flexibility.

This fact is very nicely illustrated by both Mithun’s study of lexical flexibility in Central Alaskan Yup’ik and Creissels’s study of Mandinka. Mithun (2017: 163–164), for example, considers roots for meteorological concepts, and shows that even within this small semantic domain, roots vary as to whether they exhibit flexibility. In EX the meteorological roots have predicative counterparts but in EX the meteorological roots do not.

add examples from Mithun (2017: 163)

Mithun provides similar data for the domains of clothing and instruments:

add examples from Mithun (2017: 164)

On the basis of data like these, Mithun observes, “Speakers simply know whether a given root functions as a noun and what its meaning is, and whether it functions as a verb and what its meaning is. Gaps are not predictable[.]” (Mithun 2017: 163). These gaps also vary from dialect to dialect. While the dialect in the above examples has no predicative counterpart for *taituk* ‘fog’, the Nunivak Island dialect has a pair of roots *nugu* ‘fog’ and *nungu-* ‘be foggy’.

Creissels’s (2017) study of Mandinka is another good illustration of the item-specific nature of flexibility. While Mandinka has nominal and verbal constructions that allow the predicative and referring functions of words to be distinguished unambiguously, it is not as easy to separate word stems themselves into similar classes. In Mandinka, all items are flexible, but the way in which items are flexible varies. Stems in Mandinka may be divided into three classes on the basis of their semantic behavior with regards to flexibility:

- *verbal* lexemes are those whose meaning is predictable when used to refer and therefore analyzable as a case of “morphologically unmarked nominalization”; these are always event nominalizations
- *verbo-nominal* lexemes are those whose meaning in referring constructions is idiosyncratic and therefore not predictable
- *nominal* lexemes are those whose meaning when used as predicates is predictable and limited to ‘provide someone with X’

In Mandinka, therefore, flexibility must be considered on an item-by-item basis, since the behavior of each item with regard to flexibility may differ.

In fact, flexible behavior in Mandinka is not just item-specific, but it is sense-specific as well. [Creissels \(2017: 54\)](#) reports that polysemous lexemes may show different behavior for their different senses. The word *dín*, for example, has two senses: ‘child, young (of an animal)’ and ‘fruit’. However, only the ‘fruit’ sense is available for predication; when used as an intransitive verb, *dín* may only mean ‘bear fruit’, not ‘give birth’, even though ‘give birth’ is a perfectly conceivable meaning of this word in predication. In the sense of ‘child, young (of an animal)’, *dín* behaves as a nominal lexeme, but in the sense of ‘fruit’ it behaves as a verbo-nominal lexeme. When lexical items undergo functional expansion into new discourse functions, it is also only specific senses that do so, not every sense of the word. More evidence for this comes from the diachronic development of the word *run* in English: though the word *run* when used as a predicate has numerous senses, the earliest attestations of *run* used referentially are by and large with just the sense of ‘fast pedestrian motion’ (the exceptions to this stemming from just one corpus file) ([Gries2009](#)). Other referential uses of *run* did not develop until later.

The existence of dialectal differences for flexibility as well as the unpredictable meanings of lexical items when used in various discourse functions show that the development of flexibility depends on conventionalization—whether a given form has assumed a conventionalized meaning in its role for a specific discourse function. These conventionalizations are language, dialect, item, and even sense-specific ([Croft 2000: 97](#)). Speakers can and do playfully use existing lexical items for new discourse functions, but it is not until that combination of form and discourse function is conventionalized with a specific meaning in a community of speakers that we can say the word has undergone functional expansion. An excellent illustration of this is the word *friend* in English. Prior to the public availability of the social networking site Facebook in 2006, the use of *friend* as a predicate had not been widely conventionalized. The growth of social networking then led to the very specific use of *friend* to mean ‘add as a

connection on a social networking site’. Note that it does *not* have the more general sense of ‘be a friend’ or ‘befriend’. Like with Yup’ik and Mandinka, this shows not just that flexibility is item-specific, but that the meaning of flexible uses is often item-specific as well; in many cases it is unpredictable and must be memorized by speakers.

2.3.3 Problems & critiques

Despite the robust findings in [Section 2.3.2](#), researchers have challenged the very possibility of lexical flexibility and its presence in various languages. Some of these challenges stem from the fact that certain conceptions of lexical flexibility are based on traditional ideas about the existence of large, language-specific parts of speech, and therefore subject to the same set of criticisms. Other challenges stem from precisely the facts presented in the previous section, especially that both flexibility and the meaning of flexible words are item-specific and often unpredictable, such that these words are not truly “flexible”. Moreover, languages must indicate the discourse function of their words *somehow*—this is basic to our ability to communicate. So in a certain sense, the idea that there are words which are fully ambiguous as to their discourse function is doomed at the outset. The question is really where these indications of pragmatic function live—the root, the stem, the inflected word, or the clausal context. This section summarizes the main criticisms that scholars have raised against flexible analyses. In [Section 2.4](#), we then look at alternative theories of word classes and their approach to lexical flexibility.

2.3.3.1 Methodological opportunism

A methodological problem with certain theories of flexible words is that they, like traditional theories, commit the fallacy of *methodological opportunism* (CITE: canonical references for this; borrow them from where you introduce the term above) presented in [Section 2.2.3](#). They do not apply the distributional method consistently. Instead, the criteria which separate words into categories are determined on the basis of additional theoretical commitments.

Croft (2001: §2.2.2) criticizes Hengeveld’s parts-of-speech typology on this basis, noting that Hengeveld ignores distributional evidence for classes smaller than the ones he posits in his typology (noun, contentive, etc.). Evans & Osada (2005) raise similar concerns for Hengeveld’s theory as applied to Mundari. They state that in order for two lexical items to be members of the same lexical class, they must have *equivalent combinatorics*, which is to say that their distributions should be identical (Evans & Osada 2005: 366). Evans & Osada also state that in order for a language to be flexible, that flexibility must be *exhaustive* in the sense that all members of a putatively flexible class must show equal degrees of flexibility and *bidirectional* in the sense that nouns may be used as verbs and vice versa. Both these criteria are merely different ways of reframing the broader principle that words in a class should share the same distributions (Croft 2005: 434). Evans & Osada proceed to show various ways in which these criteria are not applicable to Mundari, and that Mundari is therefore not a flexible language. At the same time, however, Evans & Osada use these facts to argue for the existence of the equally problematic categories of Noun and Verb in Mundari, using just a “canonical subset of distributional facts” (Evans & Osada 2005: 434, fn. 17). Croft’s (2005) commentary on Evans & Osada’s target article is partially devoted to critiquing them on this point. The problem of methodological opportunism is present for any analysis which assumes that languages have a small set of large lexical categories—whether that analysis is flexible or traditional.

2.3.3.2 Semantic shift

Broadly speaking, however, the primary argument against theories of flexible word classes is that they ignore a great deal of item-specific knowledge that speakers have about words and their uses in different functions (Evans & Osada 2005: §3.2; Beck 2013: 216). This issue has already been discussed in some detail in Section 2.3.2.4, but it bears explaining precisely why such item-specific knowledge constitutes a problem for theories of lexical flexibility.

For starters, when a lexical item expands into a new discourse function, there is a *semantic shift* in the direction of the meaning typically associated with the new context (Croft 1991:

74–77; 2001: 73). For example, when a property word is used in a referring expression, its meaning shifts to a person or thing with that property, not a reference to the abstract property itself. These semantic shifts cannot be attributed to some broader pragmatic principles—they are a matter of convention, and require broader uptake in a community of speakers in order to be conventionalized (as illustrated with the English word *friend* above). Because the meaning that results from this semantic shift is conventional, often idiosyncratic, and language-specific, flexible words cannot be truly productive, as is implied by the term “flexible”. There is always a conventionalized component to their meanings.

Examples of idiosyncratic and unproductive shifts in the meaning of flexible words abound in the literature. Consider agains the examples from Mundari in (3), repeated here as EX.

add Mundari examples from ex:1.3

As a predicate, the stem *buru* means ‘heap up’, but this meaning is not predictable from just the combination of the nominal sense ‘mountain’ and its predicative use. The word could have just as easily meant ‘climb a mountain’ or ‘overcome’ or simply ‘be a mountain’. No general pragmatic principles could have predicted this meaning. Likewise consider the Central Alaskan Yup’ik examples from (7c). Why does the combination of *iqeq-* ‘corner of mouth’ + *-mik* ‘thing held in one’s mouth’, ‘to put in one’s mouth’ result in *iqmik* ‘chewing tobacco’? Why not ‘oral thermometer’ or ‘toothpick’? Mithun provides many more unpredictable examples, shown in EX.

add Yup’ik examples from Mithun (2017: 165, ex. 45)

Or consider the example from Cayuga in Cayuga verbal noun, ‘it hauls logs’ = horse, repeated here as EX.

add ‘horse’ example from Cayuga

Of all the possible nominal meanings that could reasonably derive from ‘it hauls logs’—cart, tractor, ox—the fact that its nominal use means ‘horse’ is a fact specific to Cayuga that must

merely be memorized by speakers.

Conventionalizations of lexical items used in new discourse functions also vary across languages. While the principle of semantic shifts still broadly holds, the specific meanings of these conventionalizations are unpredictable. Croft exemplifies this point by comparing English *school* with Tongan *ako* ‘school / study’.

English *school* used predicatively does not mean the same thing as Tongan *ako* used predicatively, namely ‘study’. Going in the opposite direction, English *study* used referentially does not mean the same thing as Tongan *ako* used referentially, namely ‘school’. Finally, English *small* used referentially does not mean the same thing as Tongan *si’i* ‘childhood’ used referentially. (Croft 2000: 71)

Since the meanings of putatively flexible words in different discourse functions are not predictable, many scholars reason that these words cannot be truly “flexible” in the sense of multifunctional or precategoryal.

2.3.3.3 Lexical gaps

Just as unpredictable in flexible cases is which sense of a word will be coopted into the new discourse function. In Wolof, for example, the referential use of the word *ndaw* can only mean ‘young’, whereas the predicative use may mean either ‘be young’ or ‘be little, small’ (Kihm2017). We have also already seen similar lexical gaps for Central Alaskan Yup’ik in [Section 2.3.2.4](#) above. If a lexical item lacks any conventionalized use in different discourse functions, then it cannot rightly be considered flexible, or a member of a flexible word class.

2.3.3.4 Counterarguments

Pointing out that functional expansion involves both semantic shifts and functional gaps is generally intended to show that lexemes cannot be truly flexible in the sense of being multifunctional (§2.3.1.3) or precategoryal (§2.3.1.4), and that uses of the same lexical item for different discourse functions should therefore be considered cases of conversion—that is, homonymy or heterosemy. There are however two major problems with this argument.

The first is that it creates a false dichotomy between homonymy and polysemy, when in fact the two phenomena are opposite endpoints on a continuum. Debates over the lexical unity of a word—that is, whether two uses of a lexical item are homonymous or polysemous—arise from an Aristotelian desire to neatly sort those uses into distinct lexemes, when in fact reality is much more complex. If this problem sounds familiar, that is because it is exactly the same methodological problem that arises when trying to exclusively categorize words into different classes. The complex adaptive nature of language makes categorical classification at either level impossible.

As discussed in [Section 2.3.1.4](#), we know from cognitive research that mental categories are prototypal, and that the meanings of words display a polycentric, family resemblance structure. Two senses of a word are often related only tenuously through a network of intervening semantic connections or meaning chains. **Langacker1988** calls this the *network model* of category structure. [Taylor](#) points out that “[o]ne consequence of adopting the network model is that the question of whether a word is polysemous or not turns out to be incapable of receiving a definite answer.” ([Taylor 2003](#): 167)

Over time, as this lexical network expands, the meanings of a word can diverge so drastically that speakers no longer have a direct cognitive association between them. [Mithun](#) exemplifies this nicely for both Cayuga and English. Discussing morphological Verbs used as referents in Cayuga, she notes the following:

If asked the meaning of *kaqtanéhkwiḥ* [‘it hauls logs’], Cayuga speakers normally respond ‘horse’. Though it has the morphological structure of a verb, it has been lexicalized as a nominal. The literal meanings of many verbal nominals are still accessible to speakers, but the origins of others have faded, and speakers express surprise at discovering them. Similarly, when asked “What would you like for breakfast?”, most English speakers do not think about breaking their night-time fast, though they can usually be made aware of the literal meaning of *breakfast*. ([Mithun 2000](#): 413)

Lexicalization is a process and a continuum. Words can be lexicalized in new discourse functions to varying degrees. The first use of a lexical item in a new discourse function is innovative; each subsequent use then contributes further to its conventionalization.

Pointing out that functional expansion often creates idiosyncratic and unpredictable meanings essentially amounts to saying that senses of words can be highly divergent. This point is not in itself an argument against flexible analyses. Flexible words themselves may sit anywhere on the continuum from having closely connected, productive and predictable meanings, to having extremely divergent, idiosyncratic and unpredictable meanings. This is not a special fact about flexible words; it is simply true of words generally.

Croft (2001: 73) expresses concern that ignoring semantic shifts in the analysis of flexible words overlooks important insights about how such semantic shifts are patterned (specifically, the universal fact that semantic shifts are always in the direction of the word's new discourse function). Given that so many researchers have indeed ignored semantic shift when arguing for flexible analyses, Croft's concern is warranted. However, it is entirely possible to define lexical flexibility in a way that both allows for the meaning of a lexical item to encompass multiple discourse functions and acknowledges that such multifunctional uses of the word involved patterned semantic shifts. The way to do this is to ground the definition of lexical flexibility in the pragmatic functions of reference, predication, and modification rather than language-specific categories like Noun, Verb, and Adjective. I offer such a definition in [Section 2.5](#).

The second significant problem with using semantic shifts as an argument against the existence of flexible lexemes is that it proves too much. If semantic shift is taken as evidence against the lexical unity of putatively flexible words, then it must also be taken as evidence against the lexical unity of non-flexible words. Put simply, semantic shift is an analytical problem for all words, not just flexible ones.

This fact becomes clear when we ask, "What counts as a semantic shift? Just how 'large' of a change in meaning (if it were even possible to quantify such a thing) does a semantic shift require?" To illustrate this problem, consider the semantic contribution of plural marking crosslinguistically. In the canonical case, plural marking is considered inflectional rather than derivational (CITE: Corbett, Number), meaning that it does not create a new lexeme. Instead,

it modifies the meaning of the existing lexeme slightly, in line with the classic distinction between inflection vs. derivation (CITE: cite the Oxford handbooks of inflection and derivational morphology). However, there are numerous cases of words in English with more or less drastic differences in meaning between the singular and plural, and/or senses that are only available in one of the two numbers. Consider the examples in EX.

(14) English (Indo-European > Germanic)

- | | | |
|----|---------------------------------------|-----------------------------------------------------------------------------------------|
| a. | <i>air</i>
<i>airs</i> | ‘atmosphere’
‘affected manners’ |
| b. | <i>arm</i>
<i>arms</i> | ‘upper limb; anything resembling a limb’
‘weapons, firearms’ |
| c. | <i>blind</i>
<i>blinds</i> | ‘unable to see’
‘screen for a window’ |
| d. | <i>character</i>
<i>characters</i> | ‘personality, mental qualities’
‘people in a novel, play, or film’ |
| e. | <i>custom</i>
<i>customs</i> | ‘tradition; socially accepted behavior’
‘department which levies duties on imports’ |
| f. | <i>force</i>
<i>forces</i> | ‘strength, energy’
‘collection of military units’ |
| g. | <i>good</i>
<i>goods</i> | ‘excellent, high quality’
‘merchandise or possessions’ |
| h. | <i>manner</i>
<i>manners</i> | ‘way of doing something’
‘social conduct; socially acceptable conduct’ |
| i. | <i>spectacle</i>
<i>spectacles</i> | ‘visually striking performance or display’
‘pair of glasses’ |
| j. | <i>wood</i>
<i>woods</i> | ‘fibrous material in the trunk of trees or shrubs’
‘area of land covered with trees’ |

Semantic shifts for plural marking in English are not limited to just a handful of specific lexical items. Generic uses of the plural as in the expression *foxes are cunning* create a semantic shift away from a concrete entity (*a/the fox*) to a generic, unperceivable one—a use which strays from the prototypical function of nouns as concrete perceptible entities (Hopper &

Thompson 1984: 708).

As with flexible words, the semantic shifts that occur with plural marking can become so substantial that speakers no longer cognize the morphological singular and plural as members of the same word. Such is the case in the historical development of *brother* vs. *brethren* in English. *brethren* became so strongly conventionalized with its religious meaning in the plural that it was independently lexicalized as a plural-only (*plurale tantum*) noun, and the original plural underwent renewal with the emergence of the form *brothers*. This is exactly the kind of lexicalization process that occurred for many morphological verbs reanalyzed as nouns in Cayuga and many other North American languages.

A similar example comes from Chitimacha, which has a pluractional marker *-ma* indicating verbal number (plural agents, plural patients, or repeated action). In some cases the use of *-ma* is purely compositional, so that it can be considered merely an inflectional marker of verbal number. In other cases *-ma* so significantly alters the meaning of the word that it must be considered derivational. Compare the uses of *-ma* with each of the verbs in EX–EX (note that EX and EX are phrasal verbs with a preverbal particle).

(15) Chitimacha (isolate)

- | | | |
|----|--------------------|---------------------------------------------|
| a. | <i>kow-</i> | ‘call’ |
| | <i>kooma-</i> | ‘call multiple people’ |
| b. | <i>qapx cuw-</i> | ‘come back; go about’ |
| | <i>qapx cuuma-</i> | ‘travel; wander’ |
| c. | <i>qapx qiy-</i> | ‘turn together; mix, join’ |
| | <i>qapx qiima-</i> | ‘give a prayer, benediction; perform magic’ |

(Swadesh1939a)

In (15a), the use of *-ma* is entirely compositional. The presence of *-ma* indicates that the verb has a plural patient argument. In (15b), the use of *-ma* is still arguably compositional, though perhaps somewhat lexicalized given the high frequency with which the word appears in the texts. ‘travel, wander’ could reasonably be interpreted as a continued repetition of ‘go about’. In (15c), however, *qapx qiima-* has become lexicalized with a new meaning not

directly related to that of *qapx qiy-*. The diachronic connection between the two meanings is that prayers and magical incantations were traditionally accompanied by circling gestures with the arms. *qapx qiima-* originally meant ‘turn/circle around repeatedly’, but over time lexicalized with its new religious meaning in the pluractional, ‘give a prayer, benediction’. This lexicalization process parallels that of *brethren* in English. Such a range of inflectional vs. derivational uses of pluractionals is quite common crosslinguistically (CITE: Mithun on verbal number; sources on pluractionality?).

Finally, there are many languages which do not typically mark plurality on nouns (Dryer2013) (CITE: Ch. 33 of WALS), and yet have senses available in semantically plural contexts but not singular ones (where the semantic number can be understood from the clausal context, usually through verbal number marking). For example, the word *soq* in Chitimacha may mean ‘foot’ or ‘paw’ in a singular context and ‘feet’ or ‘paws’ in a plural context, but may also mean ‘tracks’ (e.g. animal tracks) in a plural context—a fairly significant and idiosyncratic shift in meaning, and one that is both language- and item-specific and thus conventional. This use constitutes a *morphologically unmarked semantic shift* in the meaning of the word, just as idiosyncratic meanings of words in cases of functional expansion also constitute morphologically unmarked semantic shifts. If we take such unmarked semantic shifts as evidence against lexical unity in the cases of flexible words, then we must also say that the ‘foot’ and ‘tracks’ meanings of *soq* constitute two distinct lexemes as well.

One might ask, if we start splitting up lexemes based on every degree of semantic shift, where does the splitting stop? This is exactly analogous to the problem of lumping vs. splitting in the context of lexical categories. The Radical Construction Grammar solution to this problem is to abandon the commitment to larger groupings of items (the major lexical categories) and acknowledge that languages consist of an interconnected network of smaller items (constructions) instead (CITE: Croft2001b). This approach has the major advantage of sidestepping unproductive debates about the existence or unity of particular lexical categories in particular languages, and shifts the focus instead to understanding the relationships and

patterns among individual constructions. This is precisely what I propose to do for lexemes as well. If we abandon the idea that all the meanings associated with a form must be in some way grouped into lexemes based on their morphosyntactic contexts of occurrence, we sidestep unproductive debates regarding homonymy vs. polysemy, and can instead focus on the relationships and patterns among the various senses associated with that form—specifically, the nature of the semantic shifts that occur between uses of the form in different discourse functions.

In sum, idiosyncratic semantic shifts are not the problem for theories of lexical flexibility that they are often taken to be. Indeed, functional expansion wouldn't be possible if hearers weren't capable of determining the meaning of a form when used in even highly unusual contexts. Innovative uses of words in new functions would be all but impossible, providing no opportunity for such innovations to receive broader adoption in the linguistic community. Each time a hearer encounters a novel use of a lexical item for the first time, they must accomplish the difficult task of discerning its meaning. This is no less true for flexible words as it is for non-flexible words, or for words whose meaning is predictable vs. unpredictable. In a certain sense, *every* use of a word is an instance of functional expansion, because every use of a word is always in a slightly different discourse and social context than the one before. The meaning of a word in a given context is highly socially and situationally dependent, and that context can change completely from one utterance to the next. Every token of a word thus necessarily appears in a new pragmatic context, and that pragmatic context slightly shapes its meaning. Language use is *is* language change. Semantic shift is therefore an integral and ubiquitous part of language use; the question is simply the degree of that semantic shift.

2.4 Functional approaches

By definition, functionalism as an approach to linguistic explanation is multifaceted. It looks to factors outside of the structural form of language as an explanation for that form—most

especially cognition, usage effects from frequency, and information structuring in discourse (Croft2001a). A functional approach to lexical categories is one that utilizes these explanatory elements. In this section I present Croft's (1991; 2000; 2001) functional theory of lexical categories, which explains crosslinguistic patterns in the coding of reference, predication, and modification as arising from the interaction between our mental categories and the needs of discourse. I then use this theory as a framework for defining lexical flexibility in Section 2.5. I begin with a brief discussion of prototype theory as it pertains to lexical categories (§2.4.1), before expounding upon typological markedness theory (§2.4.2).

2.4.1 Prototype theory

It has long been recognized that the categories of human cognition are prototypical. In a series of studies, Eleanor Rosch and colleagues demonstrate that category membership is a matter of degree, and that there are better and worse representatives of any given mental category (Rosch1973aRosch1973bRosch1975RoschMervis1975Roschetal1976Rosch1978). Prototype theory was then popularized in linguistics by Lakof1987, Langacker1987, Taylor ([1989] 2003), and Croft ([1990] 2003; 1991), among others.

The evidence for prototype effects is robust (Taylor 2003: 46–47). When asked to rate whether an item is a good example of a category, participants consistently rate prototypical members as better examples of the category than non-prototypical ones. In listing experiments where participants are asked to list members of a category, prototypical members are listed earlier and more frequently than non-prototypical members. Finally, prototypical members of a category are identified by participants as being members of the category more quickly than non-prototypical members. Each of these effects is scalar, such that individual members of a category sit anywhere on a scale of more or less prototypical.

Linguistic constructions are also subject to prototype effects (Taylor 2003: Ch. 12). HopperThompson1980, though not yet working in a prototype framework, nonetheless demonstrate that transitivity is very much a prototype category, with individual clauses showing greater or lesser degrees

of transitivity depending on their features. Ross parenthetically cites Ross 1972, Ross 1973 (CITE: both of these are cited in Taylor 2003) shows that lexical items are graded in their ability to undergo various transformations, with human beings being close to prototypical noun phrases, while inanimates, events, abstract concepts are quite atypical. Taylor (2003: §12.5) likewise points out that the transitive construction in English has steadily expanded its functions over time “to encode states of affairs which diverge increasingly from prototypical transitivity” (Taylor 2003: 235). The result of this diachronic development is significant gradation as to which verbs now lend themselves to transitivity. Taylor (2003: 236) gives the example of the transitive construction being used to imply a semantic path, in lieu of an explicit preposition.

Compare the pairs of English sentences in EX.

(16) English (Indo-European > Germanic)

Preposition

He regularly **flies across** the Atlantic.

He **swam across** the Channel.

She **swam across** our new swimming pool.

We **drove across** the Alps.

The child **crawled across** the floor.

Transitive

He regularly **flies** the Atlantic.

He **swam** the Channel.

?She **swam** our new swimming pool.

?We **drove** the Alps.

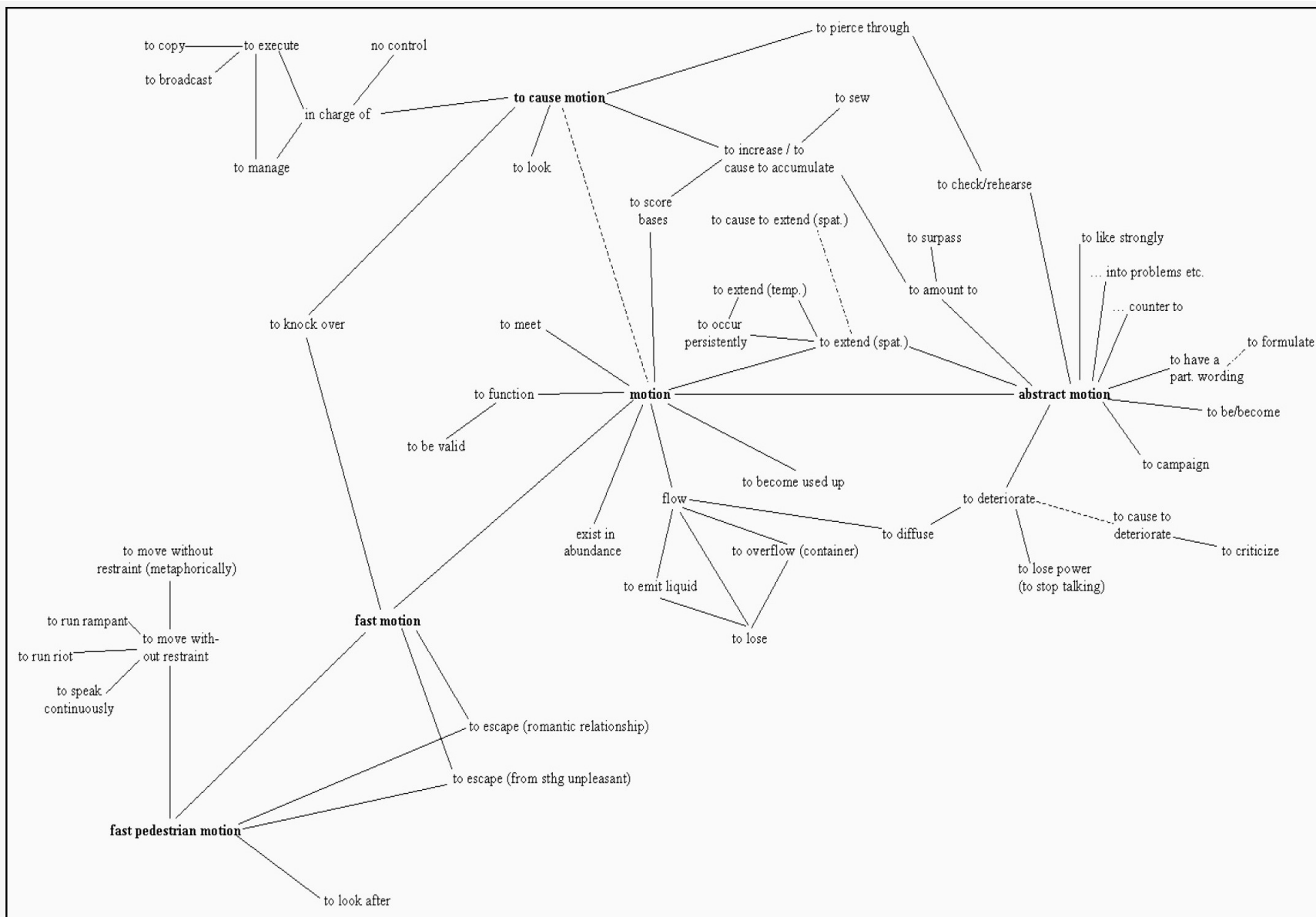
*The child **crawled** the floor.

(Taylor 2003: 236)

These examples illustrate that there are indeed better and worse members of the English Transitive Path construction.

Individual lexemes are also a type of construction, and therefore also subject to prototype effects. This is unsurprising since language forces speakers to map a non-discrete cognitive representation of the world onto discrete linguistic entities—we are forced to cut up and categorize the world around us into discrete objects and events/states so that we can refer to them and predicate statements about them. Reality, however, is not so neat. The result of this mapping is a linguistic form that imperfectly demarcates a portion of our mental world, centered on a clear prototype but with imprecise boundaries. Using a topological metaphor, we typically call some portion of our mental representation of the world a *semantic space* (Finch 2003, Dahl 2016), and that space can be graphically represented using a *semantic map*

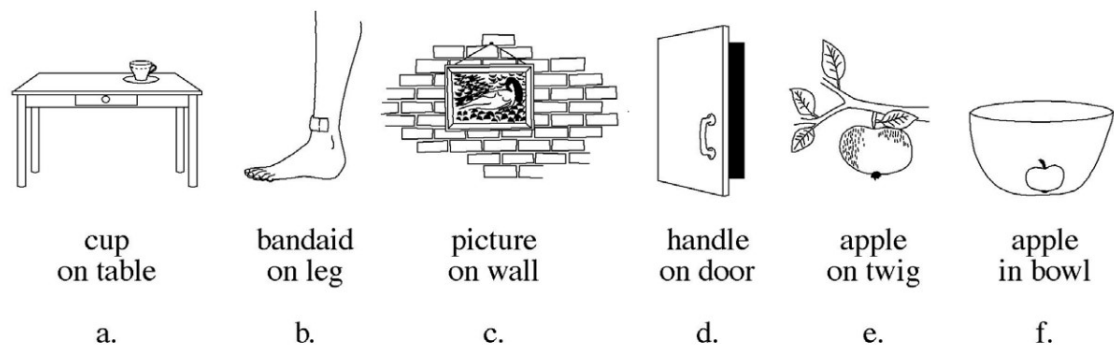
(CITE: Croft 2001 section on semantic maps; Haspelmath on semantic maps). Though semantic maps are most often used to represent a *functional* space for grammatical morphemes, they are equally applicable to lexical spaces as well. Gries2006 provides one such semantic map for the meanings of the English word *run*, shown in Figure XX, based on a comprehensive corpus analysis. BowermanChoi2001 report on earlier work by BowermanPederson1991 in which they construct a semantic map of spatial relations on the basis of data from 38 languages (25 families), with a relation indicating prototypical support from below (ON) at one end and a relation indicating prototypical containment (IN) at the other. As pictured in Figure XX, words in different languages cut up this semantic space in different ways.

Figure 2.5: Semantic map of English *run* (Gries2006)

These examples illustrate that word meanings are polycentric and cover a range of possible uses, as mentioned in [Section 2.3.1.4](#). Some of these uses may be more prototypical than others. The expression *apple on a twig* is a slightly less prototypical use of *on* than *apple on a table*. The fact that words cover a range of uses, and that some of these uses are more prototypical than others, is an important component of the typological markedness theory of lexical categories.

Even the formal categories that linguists use to describe linguistic structure tend to be prototypical ([Taylor 2003](#): xii, 201) (CITE: Categories and linguists' categorizations, in LT). [Taylor \(2003: §11.1\)](#) argues that linguists' conceptions of the formal labels *word*, *affix*, and *clitic* are prototypical in nature, with better and worse members of the category. [Haspelmath 2005](#) likewise shows that simple structural definitions of these categories are inadequate, and reframes the word–affix continuum in functional terms instead. Much research in the Canonical Typology framework (CITE: add citations from your LT article) also demonstrates the prototypical nature of linguists' categories. Though Corbett is careful to distinguish between a *canon* and a *prototype* / *exemplar* (CITE: cite a place where Corbett makes this point), his accumulated work nonetheless shows that linguists view phenomena in the world's languages as better or worse instances of various descriptive categories.

What type of category are lexical categories then? Are word classes categories of human cognition, categories within particular languages, categories of languages generally, or analytic categories of linguists? Or some combination of these? Typological markedness theory posits that parts of speech like noun, verb, and adjective are not categories of particular languages. Languages have constructions, not parts of speech. Speakers, however, have *mental prototypes* of objects, actions, and properties. And although there is no *singular* Noun construction in English that would correspond to the mental category of OBJECT, there are numerous constructions in English which have the function of indicating **reference to an object**, such as the Definite Article construction or the Transitive Subject construction. Likewise, there is no one construction—in English or any language—that can be definitively called the



1. English ON IN

2. Japanese UE NAKA

3. Dutch OP AAN IN

4. Berber X DI

5. Spanish EN

Figure 2.6: Crosslinguistic differences in the encoding of spatial relationships (BowermanChoi2001)

Verb construction or the Adjective construction, but there are plenty of constructions which have the function of predicating or attributing properties. Naturally, then, speakers are more likely to use referring constructions when talking about something which they mentally categorize as an object, predicating constructions when talking about something they conceive of as an action, and modifying constructions when talking about something they conceptualize as a property.

Speakers' conceptualizations, however, are fluid. Speakers often conceptualize things in non-prototypical ways. They may construe events as bounded entities that they can refer to, or objects as properties with duration. As a result, speakers often use words in constructions that do not align particularly well with the word's meaning, such as the appearance of an action word like *sing* in a referring construction like the Gerund in the phrase *his singing was beautiful*. When speakers use words in this atypical manner, those words are much more likely to be marked in some way—whether morphologically, behaviorally, semantically, or frequently (CITE: Croft). As a consequence of this tendency, clear asymmetries emerge between the prototypical vs. non-prototypical uses of object words, action words, and property words. It is the unmarked use of these words that most closely aligns with linguists' traditional conceptions of noun, verb, and adjective. Parts of speech as traditionally conceived are nothing more than the emergent effects of our cognitive prototypes on language. They do not have any real status in grammar or individual grammars. This is the fundamental idea behind typological markedness theory. Section 2.4.2 lays out this theory in more detail.

A last clarifying point is in order. Recognizing the existence of prototype-based categories, many linguists have described parts of speech as prototypical. Dixon & Aikhenvald (2004: 1–2), for example, says that the word classes noun, verb, and adjective each have a “prototypical conceptual basis” and “prototypical grammatical functions”. Taylor (2003: 217) states, “A prototype view of NOUN entails that some nouns are better examples of the category, while others have a more marginal status.” But languages have constructions, not parts of speech, and individual constructions are not gradient (CITE: Croft, Beyond Aristotle and gradience).

What linguists are in fact observing when they say that parts of speech are prototypal is not gradation in *linguistic categories* like noun, verb, and adjective (since those don't exist), but rather gradation in the *mental categories* of objects, actions, and properties, which do indeed exhibit prototype structures, and which therefore have emergent effects on the organization of constructions in languages.

2.4.2 Typological markedness theory

I have already previewed various aspects of typological markedness theory at different points in this thesis. In this section I present a concise overview of the specific claims made by this theory, and some of the evidence for those claims. The phrase *typological markedness* or *typological markedness asymmetries* simply refers to an implicational universal regarding the behavior of basic versus non-basic members of a conceptual category. At its simplest, the theory posits that less basic or prototypical members of a category are marked in some way; basic or prototypical category members are unmarked by comparison (Greenberg1966). This *cognitive* markedness is then realized *linguistically* in a number of ways. The marked member of a category *may* be literally marked with some kind of affix or other overt morphological indicator, but this is just one of the ways an item can be a marked member of a category. The marked member of a category may also be less frequent, or have a smaller range of inflectional / distributional possibilities, or be semantically more complex. It is important to emphasize that *typological* markedness does *not* always entail *formal* markedness. Typological markedness is an *implicational* universal rather than an *absolute* universal. The more marked members of a category must be *at least as marked* as the unmarked member, but this does not preclude the possibility of all members being *equally* marked. Formal markedness is merely an emergent tendency of structures to reflect cognitive markedness.

As applied to word classes, typological markedness theory states that the most unmarked discourse functions for object, action, and property words are reference, predication, and modification, respectively. Therefore, when a lexical item is used for a function that does

not align with its prototypical meaning, typological markedness theory predicts that it will be marked. Again, it must be emphasized that not *every* instance of a lexical item being used in a non-prototypical function will be marked in comparison to its prototypical function; but it will always be *at least as* marked. This theory of typological markedness for the major discourse functions is laid out in detail by Croft in various publications (Croft 1991; 2000; Croft 2001; Croft & Lieber 2012). It is also important to understand that typological markedness theory is *not* a theory of parts of speech in the sense of large partitionings of the lexicon into categories like noun, verb, and adjective. Instead, noun, verb, and adjective are epiphenomenal, crosslinguistic markedness patterns that arise from the interaction of semantic prototypes (object, action, property) and their use in different discourse functions (reference, predication, and modification). They are not categories of particular languages.

Throughout this thesis, I have used the term *discourse function* to refer to the functions of reference, predication, or modification. These are what Croft (1991: 51) calls *pragmatic functions* or *propositional act functions* following the tradition of pragmatics and speech act theory in philosophy (Austin 1962; Searle 1976). These three functions are taken as fundamental to human communication, arising out of the communicative intent behind what speakers are actually attempting to *do* with language. This perspective was articulated early on by Sapir:

There must be something to talk about and something must be said about this subject of discourse once it is selected. This distinction is of such fundamental importance that the vast majority of languages have emphasized it by creating some sort of formal barrier between the two terms of the proposition. (Sapir 1921)

A similar point is made by Croft while articulating his theory of typological markedness as applied to lexical categories: “[n]o matter how complex a given situation is in terms of the number of entities involved and the number and kinds of relations that hold between them, a human being attempting to describe it in natural language must split it into a series of reference-predication pairs[.]” (Croft 1991: 124)

Modification is generally seen as less central a function than reference and predication, as illustrated by its lack of mention in the quotes above. For example, Hengeveld1992 takes the reference-predication dichotomy to be fundamental, yielding the major categories of noun and verb, while the modification function then combines with these two functions to yield the major categories of adjective and adverb, respectively. The primacy of the reference-predication distinction also appears to be reflected structurally in the world's languages, which do not always have dedicated means for encoding modification, but appear to always have constructions dedicated to reference and predication.

Croft (1991: 123) defines the pragmatic functions in terms of their discourse functions, following work in the discourse-functional tradition (Chafe1976Givon1981Hopper & Thompson 1984Chafe1987DuBois1987) (CITE: all cited in Croft 1991). Previous research defines *referents* as “discourse-manipulable participants” (DuBois1980; Givon1981; Hopper & Thompson 1984: 711; Kibrik2011) (CITE: DuBois and Givon are both cited in HopperThompson1984), *predicates* as reported events (Hopper & Thompson 1984: 726), and *modifiers* as a mix of these two functions (Thompson 1989). Croft123 puts forth the following definitions instead:

- the act of *reference* identifies a referent and establishes a cognitive file for that referent
- the act of *predication* ascribes something to a referent
- the act of *modification* enriches the cognitive image of the referent with an additional feature

The particular pragmatic function chosen for any given mention of a concept is then just a matter of how the speaker chooses to portray or construe that particular concept—whether as a referent, predicate, or modifier (Croft 1991: 100); as CroftLier2012 note, “apparent instances of ‘fuzziness’ are actually variable construals” (CroftLier2012).

With this understanding of discourse functions in mind, we can restate the thesis of typological markedness theory as applied to lexical categories: Noun, verb, and adjective

are epiphenomenal markedness patterns that arise from the use of different semantic prototypes (objects, actions, and properties) in different discourse functions (reference, predication, modification). Uses of these semantic classes in non-prototypical functions are typologically marked. As mentioned, there are four ways in non-prototypical uses can be marked: structurally, behaviorally, semantically, and/or frequently.

The first type of marking, *structural coding* or *formal marking*, refers to the fact that non-prototypical uses of lexical items are at least as formally marked as prototypical ones. Structural coding in this context refers specifically to “dedicated formal markers in a specific language that indicate a lexeme’s syntactic function” (Croft & van Lier 2012). Figure XX is a schematic representation of some of the formal realizations of these markedness patterns. It indicates the different morphosyntactic means that languages tend to develop for marking each of the non-prototypical uses of words. For instance, participle constructions are one way that languages have of indicating the non-prototypical case of an action word being used for modification.

combine Croft (2000: 89) with Croft & van Lier (2012: 62)

The second way in which non-prototypical uses of words can be marked is in terms of their *behavioral potential*, that is, the range of combinatorial possibilities for that lexical item. This is most clearly illustrated with an example from inflection: in many languages, property words used in predicate constructions are limited in their inflectional possibilities. In Munya, for example, property words functioning as predicates cannot inflect for person and number of the subject, and cannot take the imperfective marker, perfective marker, or direct evidential marker (Bai 2019). The only grammatical markers allowed in property predication clauses are the stative aspect marker, a clause-final particle, and an egophoric marker. Hopper & Thompson’s (1984) study of the discourse functions of different parts of speech is largely a study of behavioral potential. They conclude that “the closer a form is to signaling this prime [prototypical] function, the more the language tends to recognize its function through morphemes typical of the category—e.g. deictic markers for [Nouns], tense markers for [Verbs].”

(Hopper & Thompson 1984: 703, abstract). Croft advances a cognitive explanation for these behavioral markedness patterns:

In general, only the core members of the syntactic category will display the full grammatical behavior characteristic of their category because only they have all the semantic characteristics that the characteristic inflections tap into. This is to say that the inflectional categories of the major syntactic categories have been “tailored” to their semantically core members. This is an example of a processing constraint: languages inflect only for those properties that are of relevance to core members of the category; they do not inflect for properties of peripheral members of the category that are not of relevance to the core members of the category. (Croft 1991: 86)

Non-prototypical uses of words may also be marked semantically by a *semantic shift* in their meaning towards the semantic class prototypically associated with the discourse function they are found in (Croft 2000: 96; Croft 2001; Croft & Lier 2012). I have already discussed the semantic shifts that occur in functional expansion in some detail in Section 2.3.3.2. Croft (1991: 60–61) makes the even stronger claim that non-prototypical uses of words will *always* be marked semantically, making semantic markedness an absolute rather than implicational universal.

These semantic shifts are caused by a combination of conventionalization and *coercion*, wherein the meaning of the construction is imposed on the meaning of the lexical item (Pustejovsky 1991: [69, 108]; Croft 1991: [252]; Panther & Thornburg 2007; (CITE: Oxford handbook of cognitive linguistics)). For example, predicate nominals (wherein an object word is used in a predicate construction) involve coercion from denoting an object to denoting classifying or equational relations (Croft 1991: 69). In Nuuchahnulth, for example, nominal predicates are always semantically durative and interpreted as either existential, classifying, or identifying expressions (Nakayama 2003).

The final way in which words used in atypical functions may be marked is in terms of their frequency. Croft (1991: 59, 87) also refers to this as *textual markedness*. Frequential markedness predicts that lexical items are used more frequently in their prototypical functions than in non-prototypical ones. This means that object words should be most frequent in their

use in referring constructions, and that referring constructions should most frequently denote objects (Croft 1991: 87).

The field of linguistics has accumulated a good deal of empirical evidence in support of the typological markedness theory of lexical categories. Croft (1991) provides empirical evidence from 12 languages for each of these markedness patterns. Dixon 1977 also provides evidence of typological markedness patterns as they relate to property words, using a combination of structural and behavioral evidence. As mentioned, Hopper & Thompson's (1984) study also provides empirical support from a variety of languages for markedness in terms of behavioral potential. Stassen (1997) is a massive study of intransitive predication in 410 languages, demonstrating the marked behavior of non-action words when used in predicate constructions.

Having explicated the basic tenets of typological markedness theory, I now turn to re-framing the concept of lexical flexibility in a way that utilizes this framework.

2.5 Lexical flexibility: A functional definition

Within the framework of typological markedness asymmetries, lexical flexibility can be understood as follows:

lexical flexibility The use of a lexical item (root, stem, or inflected word) in more than one discourse function (reference, predication, or modification) with no difference in structural coding.

This definition qualifies as a valid *comparative concept* in the sense of Haspelmath 2010, because it is couched in terms of universal *functions* rather than language-specific *structures* (Croft 2016). It also has the advantage of being intentionally equivocal with respect to the lexical and cognitive unity of the item, and with respect to the linguistic level (root, stem, or inflected word) at which the flexibility is realized. In some cases when a single lexical form appears in more than one discourse function, speakers may have a close cognitive association

between the two uses. This is most likely the case for the predicative and referential uses of the word *run* in the phrases *I run every morning* and *I'm going for a run* respectively. In other cases, speakers may have little to no awareness of the diachronic connection between uses of a form. For example, the use of *run* in the sense of *to run a print job* is extremely distant from the prototypical “fast pedestrian motion” sense in the semantic network for that form (Gries2006; see also Figure XX). It is unlikely that these two senses are closely cognitively connected by most speakers, despite the fact that they both share a predicating function. The definition of lexical flexibility given here allows for any degree of semantic shift. Croft2001 admits of this possibility explicitly: “It of course a priori possible to construct a typological classification of parts-of-speech systems using only structural coding and allowing any degree of semantic shift.” (Croft2001) Of course, I am not concerned here with constructing a classification of parts of speech—quite the opposite, in fact. This definition of lexical flexibility is intended to delimit precisely the cases where a language does *not* provide formal indicators of discourse function.

Allowing for any degree of semantic shift does *not* imply that semantic shift is in any way unimportant for understanding lexical flexibility. On the contrary, semantic shift is a key component of the process of functional expansion that leads to flexibility in the first place (see below). Moreover, carefully circumscribing the phenomenon of lexical flexibility without regard to the degree and type of semantic shifts involved puts us in a position to then compare the semantic shifts that occur in cases of lexical flexibility with those that occur in cases of overt derivation. This raises the intriguing question of whether semantic shifts in flexible cases differ in principled ways from overt derivation. Mithun (2017: ???) shows that for Central Alaskan Yup'ik the types of semantic relationships between flexible uses of words mirror those between basic and derived words. This would suggest that functional expansion follows the same principles as overt derivation. However, much more research is needed in this area.

As we have seen, a great abundance of evidence also shows that the meaning of any

given combination of form and discourse function is a matter of convention, and often highly idiosyncratic (§??sec:2.3.2.4); §2.3.3.2). This fact suggests that flexible items are not truly “flexible” in the sense that speakers can use any lexical item for any discourse function and expect hearers to be able to infer their meaning from context. We know that item-specific gaps in usage exist. Certainly, novel cases of forms being used in new discourse functions do occur, or else it would not be possible for functional shift to happen in the first place. But these cases are necessarily restricted by the cognitive limits on our ability to deal with ambiguity. If it were truly the case that any lexical item could be used in any discourse function at any time, it would barely be possible for hearers to interpret the intended pragmatic effects of each word. Instead, flexibility must rely on a degree of *conventionalization*. Conventionalization in turn implies *time*—conventionalization is a diachronic process. Thus, *lexical flexibility is best understood as a synchronic pattern resulting from the diachronic process of functional expansion*, where functional expansion is defined as follows:

functional expansion A diachronic expansion in the use of a lexical item (root, stem, or inflected word) into a new discourse function (reference, predication, or modification) with no additional structural coding.

Cases of lexical flexibility therefore arise whenever a new combination of form and discourse function is conventionalized in a community of speakers. This understanding of lexical flexibility is in line with cognitive research on lexicalization and constructional change. Functional expansion occurs because of speakers’ need to construe concepts in different ways—as objects, actions, or properties. The semantic shifts that occur during functional expansion are then the result of coercion by the new constructional context, and the resultant meaning then becomes conventionalized as the meaning of that particular form in that particular discourse function (Croft 1991: 108).

If lexical flexibility is the result of a diachronic process, it should be possible to enumerate some of the specific pathways which give rise to it. Here I will mention just a few. One pathway is insubordination, whereby subordinate clauses in a language are reanalyzed as

main clauses (CITE: Mithun, Evans; can probably use the same citations as earlier in this chapter) (see also §2.3.2.2). Insubordination frequently results in formal similarities between noun phrases and verb phrases, and this formal ambiguity can abet the functional expansion of lexical items from referential to predicative uses and vice versa.

A second pathway to lexical flexibility is relexicalization (or more precisely, reconventionalization). This is the process that occurred in the case of morphological verbs being reanalyzed as nouns in many North American languages (see §2.3.2.3) and certain English plurals like *brethren* or *arms*. In these cases, the conventionalized meaning associated with the form changed (e.g. from ‘it hauls logs’ to ‘horse’), so that the meaning more closely reflected its new discourse context.

A third pathway is topicalization, exemplified in the Wakashan family. Jacobsen, Jr. (1979: 122, 142) observes the formal similarity between the Definite Article and the Third Person Singular Indicative markers in Wakashan languages, and argues for a diachronic connection between the two. It is likely that cleft constructions such as ‘it was the dog that ran’ became so common that speakers started to reanalyze the topicalized cleft as a definite noun phrase, ‘the dog’, thereby creating a formal similarity between referring expressions and predicating expressions.

Each of these pathways results in the functional expansion of lexical items into new discourse contexts with no new overt structural coding. Of course, functional expansion can also occur without any other accompanying grammatical changes. This happens in any instance where speakers simply use stems in new discourse functions. Lexical flexibility is the natural and expected result of the fact that non-prototypical uses of words are *not* always structurally marked—as allowed for by the fact that typological markedness is implicational and not absolute—even if they are marked in other ways. The use of additional structural coding in cases of functional expansion is not obligatory, but merely a statistical tendency. Lexical flexibility occupies the theoretical space where structural coding asymmetries fail to apply.

When viewed in this light, *lexical flexibility is not so much of a problem as it is a design feature of language*. The presence of lexical flexibility should be *expected* in every language, not treated as exotic. The cognitive-typological approach outlined in this chapter inverts the lexical flexibility question: the interesting question is not why some languages fail to make distinctions in parts of speech (thereby framing lexical flexibility as a *deficit*), but rather why languages develop specialized constructions for different discourse functions in the first place. Lexical flexibility exists in any area of the grammar where specialized function-indicating morphology has yet to develop. Flexibility should therefore be considered the default state of affairs for language. Gil (CITE: both Gil publications on IMA languages) has in fact argued, partially on the basis of data from highly flexible languages, that early human language must have been *isolating-monocategorical-associational* before the development of dedicated function-indicating morphology.

The idea that the “natural state” of language is monocategorical or acategorical would seem to conflict with the point made above that lexical flexibility can result from diachronic processes, but the two positions are not mutually exclusive. Languages develop constructions which indicate different discourse functions, but languages are also subject to counteracting pressures. This is a classic case of competing motivations: on the one hand, the frequency with which speakers need to perform the discourse functions of reference, predication, and modification all but ensures the development of constructions dedicated to indicating those functions; on the other hand, speakers need to construe states of affairs in various ways—as objects, actions, or properties—creating pressures which have the potential to level those formally marked distinctions. Reconventionalization and the reanalysis of cleft constructions could both be viewed as diachronic processes motivated by this latter pressure.

In sum, then, lexical flexibility is a natural result of the cognitive and diachronic forces at work in language. Defining lexical flexibility in terms of typological markedness (or more accurately, the lack of formal marking for otherwise marked uses) provides a crosslinguistically applicable definition of the phenomenon which avoids methodological opportunism

while still recognizing that lexical flexibility requires some degree of semantic shift and conventionalization. With this definition in place, the remainder of this thesis turns to exploring the prevalence of lexical flexibility in English and Nuuchahnulth and the semantic behavior of words in cases of functional expansion.

To Do

■ citation	22
■ citation	23
■ citation	29
■ add distributional tables	31
■ citation	31
■ citation	33
■ add a section on mixed / hybrid categories (including Lois, Malouf, Nikolaeva)	33
■ add examples of flexible vs. rigid languages, with explanations	35
■ citation	35
■ cross reference	37
■ cross reference	37
■ cross reference	37
■ cross reference	37
■ add examples of multifunctionality	37
■ cross reference	37
■ add examples of conversion	37
■ add lead-out	37
■ cross reference	37
■ cross reference	37
■ citation	37
■ citation	38
■ add examples from Kihm (2017: 87)	39
■ add diagram for monocentric vs. polycentric categories	40
■ citation	41
■ add parts-of-speech tables	44
■ citation	44
■ citation	45
■ citation	45
■ citation	45
■ cross reference	47
■ cross reference	47
■ add examples of from Mithun (2017: 167–168)	47
■ add examples of word-level flexibility	47
■ cross reference	48
■ cross reference	48

add examples from Mithun (2017: 163)	48
add examples from Mithun (2017: 164)	48
citation	51
cross reference	52
add Mundari examples from ex:1.3	52
cross reference	53
add Yup'ik examples from Mithun (2017: 165, ex. 45)	53
cross reference	53
cross reference	53
add 'horse' example from Cayuga	53
citation	56
citation	56
cross reference	56
cross reference	58
cross reference	58
cross reference	58
cross reference	58
citation	58
citation	58
citation	59
citation	61
cross reference	61
citation	62
cross reference	62
cross reference	62
citation	64
citation	64
citation	64
citation	66
citation	66
citation	69
citation	69
cross reference	70
combine Croft (2000: 89) with Croft & van Lier (2012: 62)	70
citation	71
cross reference	73
citation	75
citation	76