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 summarize approaches to lexical categories from several functionalist perspectives—cognitive linguistics, typology, and construction grammar. I conclude by offering a revised formulation of lexical flexibility which is more in line with this functional research.

Draft

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 1980s, many linguists came to view categories as prototypal, 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 > prototypal > 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 Universalist

Historically and still presently, many researchers assumed that a small set of lexical categories are basic and universal to all languages (**BolingerSears1981**Croft 1991: 2**Payne1997**; 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 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	*Auraicept na n-Éces* 'The scholars' primer'	Longarad
Occitan	1327	*Leys d'amors* 'Laws of love'	Guilhèm Molinièr
Welsh	1382–1410	*Llyfr Coch Hergest* 'Red book of Hergest'	unknown
Tuscan	1437–1441	*Grammatica della lingua toscana* 'Grammar of the Tuscan language'	Leon Battista Alberti

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

Language	Year	Title	Author
Castilian	1492	*Gramática de la lengua castellana* 'Grammar of the Castilian language'	Antonio de Nebrija
French	1530	*L'Éclaircissement de la langue francoyse* 'Explication of the French language'	John Palsgrave
German	1534	*Ein Teutsche Grammatica* 'A German grammar'	Valentin Ickelsamer
Basque	1545	*Linguæ Vasconum Primitiæ* 'First fruits of the Basque language'	Bernard Etxepare
Totonac	1539–1554	*Arte de la lengua totonaca* 'Grammar of the Totonac language'	Andrés de Olmos
Nahuatl	1547	*Arte para aprender la lengua mexicana* 'Grammar for learning the Mexican language'	Andrés de Olmos
Tarascan	1558	*Arte de la lengua tarasca de Michoacán* 'Grammar of the Tarascan language of Michoacán'	Maturino Gilberti
Dutch	1559	*Den schat der Duytsscher Talen* 'The treasure of the Dutch language'	John III van de Werve
Quechua	1560	*Grammatica o arte de la lengua general de los Indios de los Reynos del Peru* 'Grammar or Art of the General Language of the Indians of the Royalty of Peru'	Domingo de Santo Tomás
Tzeltal Maya	1571	*Ars Tzeldaica* 'Tzeltal Grammar'	Fray Domingo de Hara
Zapotec	1578	*Arte en lengua Zapoteca* 'Grammar in the Zapotec language'	Juan de Córdova
English	1586	*Pamphlet for Grammar*	William Bullokar
Mixtec	1593	*Arte de lengua Mixteca* 'Grammar of the Mixtec language'	Antonio de los Reyes
Timucua	1614	*Gramatica de la lengua Timuquana de Florida* 'Grammar of the Timucua language of Florida'	Francisco Pareja
Narragansett	1643	*A key into the language of America*	Roger Williams

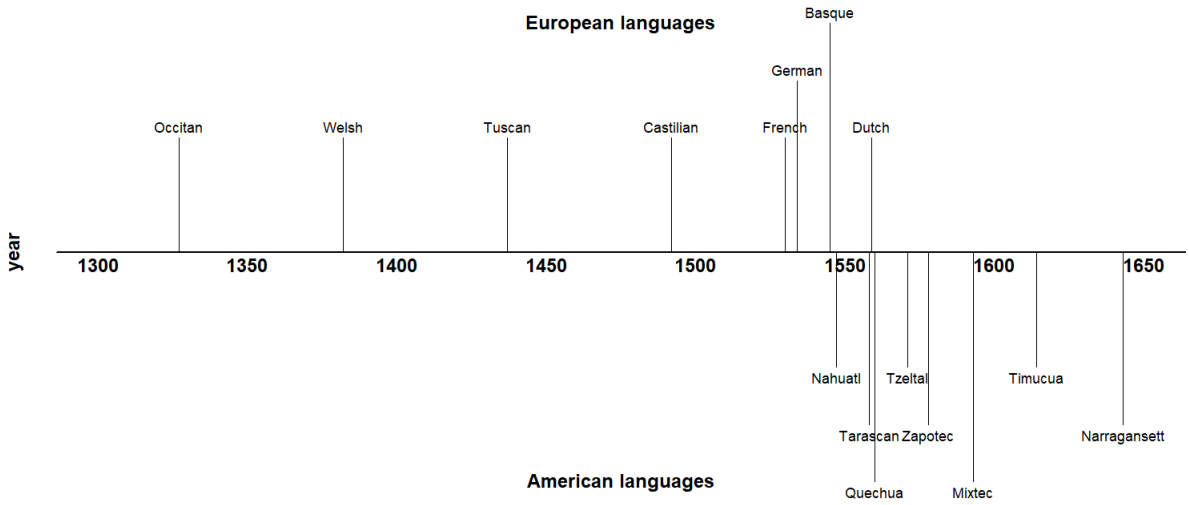

Figure 2.1: Approximate date of some of the first grammatical descriptions of European vs. American languages

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 suspectible 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, categorial 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 Culicover [**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**]).

2.2.2 Relativist

American ethnographers in the tradition of Franz Boas questioned the univeralist assumption in a programmatic and comprehensive way. Writing on grammatical categories, Boas 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**). Boas also strongly influenced Leonard Bloomfield, who treated language as a scientific object and, in applying Boasian relativism, 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**).

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 Verb (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**Haspelmath2010LaPolla2016**) (CITE: You

could probably cite your 2013 article for this too.). This position is discussed further in §2.4.

2.2.3 Structuralist

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**). 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**)[11]Croft2001. 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) a. Jack is cold.
b. * Jack colds.

(12) a. Jack is happy.
b. * Jack happies.

(13) a. * Jack is dance.
b. Jack dances.

(14) a. * Jack is sing.
b. Jack sings.

We can see that *cold* and *happy* have the same distribution (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 this distributional test 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: *cold*, *happy*, etc.	✓	✘
Verb: *sing*, *dance*, etc.	✘	✓

2.3 Flexible approaches

2.4 Functional approaches

2.5 Lexical flexibility: A functional definition

To Do