

## Oxford Handbooks Online

### The Morphology of Words

James P. Blevins

The Oxford Handbook of Language Production

*Edited by Matthew Goldrick, Victor S. Ferreira, and Michele Miozzo*

Print Publication Date: Apr 2014

Subject: Psychology, Personality and Social Psychology, Cognitive Psychology

Online Publication Date: Jul 2014

DOI: 10.1093/oxfordhb/9780199735471.013.018

### **[–] Abstract and Keywords**

Morphology is usually understood as the branch of linguistics that investigates word structure, a topic of central relevance to the systematic study of language and language processing. The Western grammatical tradition begins with the identification of words as the smallest meaningful elements of speech, a conception that survives largely intact in contemporary word-based models of morphology and grammar. Synchronic, historical, and behavioral evidence also suggests that words are not only organized into syntagmatic units but also into paradigmatic collections. On the syntagmatic dimension, words are composed of morphs and themselves form parts of larger syntactic constructions. Orthogonal to these structures, inflected and derivational forms exhibit an organization into inflectional paradigms and larger morphological families. This chapter outlines some of the linguistic issues that arise in describing words and their structure.

Keywords: morphology, morphological theory

---

### **Why words?**

Morphology is usually understood as the branch of linguistics that investigates “word structure.” One dimension of this investigation focusses on “the study of morphemes and their arrangement into words” (Nida, 1949, p. 1). A second, complementary, dimension examines “the forms of words in different uses and constructions” (Matthews, 1991, 3). Words are central to both of these investigations, reflecting the position they have occupied in models of language description and analysis since classical antiquity. The Western grammatical tradition begins with the identification of words as the smallest meaningful elements of speech, a conception that survives largely intact in contemporary word-based models of morphology and grammar. A similar perspective is represented in reference and pedagogical traditions that describe the lexical resources of a language primarily in terms of its word stock. Even Bloomfield (1933) freely conceded the practical value of words for dictionaries and reference grammars.

For the purposes of ordinary life, the word is the smallest unit of speech. Our dictionaries list the words of a language; for all purposes except the systematic study of language, this procedure is doubtless more useful than would be a list of morphemes. (p. 178)

The treatment of words as the smallest meaningful units of language is also implicit in procedures of phonemic analysis, which distinguish allophones from phonemes in terms of whether a given phone distinguishes the meaning of two words. The salience of words reinforces their role as the primary exponent of form variation in a language. Words are the smallest grammatical units that are demarcated with any consistency in the speech stream. They are also the smallest units that can function as utterances in isolation. Stages of first language acquisition are commonly defined in terms of the number of words per utterance. Behavioral correlates of word frequency, measured both in terms of “types” and “tokens,” have been observed in psycholinguistic and

neurolinguistic studies of language processing. Synchronic, historical, and behavioral evidence also suggests that words are not only organized into syntagmatic units but also into paradigmatic collections. On the syntagmatic dimension, words are composed of morphs and themselves form parts of larger syntactic constructions. Orthogonal to these structures, inflected and derivational forms exhibit organization into inflectional paradigms and larger morphological families.

Word structure is thus of central relevance to the “systematic study of language” and language processing. To set the context for the experimental and computational studies presented in subsequent sections, this chapter outlines some of the linguistic issues that arise in describing words and their structure.

## Words

A treatment of words and word structure must begin by acknowledging that the definition of words is far from settled. Despite this lack of consensus, there are points of broad agreement. Most contemporary accounts recognize the usefulness of distinguishing three types of units, each corresponding to an established sense of the notion “word.” Phonological words or **WORDFORMS** are simple sequences of phonemes (or graphemes). Grammatical words are phonological words with a morphosyntactic interpretation. Lexical words or **LEXEMES** correspond to the units that would be entered in a dictionary. Each of these units has a characteristic role in the morphology of a language, and each determines a corresponding notion of “word count.” Yet the partial overlap between these notions also raises descriptive and analytic issues that need to be clarified.

## Grammatical and Phonological Words

The notion of “word” proposed in Bloomfield (1926) provides the classic definition of the “grammatical” word. Bloomfield’s notion of a “minimum free form” also comes closest to capturing the sense of word as it is normally understood.

A minimum free form is a **WORD**.

A word is thus a form which may be uttered alone (with meaning) but cannot be analyzed into parts that may (all of them) be uttered alone (with meaning). (p. 156)

On this interpretation of “word,” the three homophonous occurrences of *cut* in (1) are different words, realizing the preterite, past participle, and infinitive forms of the verb *CUT*. This is the notion of word that is most relevant to measures of token frequency, which treat each occurrence of a form as a separate “token.”

### (1)

- a. He *cut* the flowers. (PRETERITE)
- b. He has *cut* the flowers. (PAST PARTICIPLE)
- c. He will *cut* the flowers. (INFINITIVE)

The Bloomfieldian notion of a “minimum free form (with meaning)” also captures the sense of “word” that is mainly relevant for (extended) “word and paradigm” models (Robins, 1959; Anderson, 1992) and for other “word-based” approaches (Blevins, 2006). As Robins (1959, p. 120) puts it, the word in this sense is “a grammatical abstraction,” but it is a maximally useful abstraction.

The word is a more stable and solid focus of grammatical relations than the component morpheme by itself. Put another way, grammatical statements are abstractions, but they are more profitably abstracted from words as wholes than from individual morphemes. (p. 128)

In most languages, the abstraction of words is facilitated by cues that enhance their perceptual salience. Open-class items are often subject to a minimum word constraint, whether measured in terms of moras, syllables, or metrical feet, and there is experimental evidence that speakers exploit these constraints in the segmentation of continuous speech (Norris, McQueen, Cutler, & Butterfield, 1997). Words often define the positions at which stress, pitch, or other suprasegmental features are realized, and word edges maybe marked by processes, such as boundary lengthening (Bybee, 2001). The perceptual salience of words is further enhanced by the fact that words (unlike subword units, such as phonemes or morphemes) may stand on their own as independent utterances. In

addition, if there is any basis to notions like “the one-word stage” (Dromi, 1987), it would appear that the word is the canonical utterance during early stages of language acquisition. As one might expect, the functional load of individual cues varies across languages, reflecting general differences in phonological systems, so that no single cue identifies words cross-linguistically. This leads, as Robins (1959) acknowledges, to discrepancies between the notion of “word” relevant for the description of grammatical relations, and the sense of “word” that is marked phonetically.

### Wordforms

The qualification “grammatical” reflects this contrast between the Bloomfieldian word and “phonological” (or, by extension, “orthographic”) words. A more concise term for minimal free forms WITHOUT a fixed meaning or function is “wordform.” In example (1) above, the phonological word *cut* realizes the preterite, past participle, and infinitive forms of the verb *CUT*. The same wordform realizes other forms of *CUT*, as well as the corresponding noun in *he has a cut*.

Wordforms are, by definition, the units that are most clearly demarcated in the speech stream by phonetic cues. As sequences of phonemes (or graphemes), they are also the most straightforward units to distinguish and count. Although wordforms are not traditionally of significant grammatical interest, they are central to many current frequency-based models. In accounts that treat phonetic reduction as a frequency effect, it is often wordform frequency that is taken to determine the rate or extent of reduction. Hence in the frequency-based model of Bybee (2010), the operative notion of “words” corresponds to “wordforms.”

A robust finding that has emerged recently in quantitative studies of phonetic reduction is that high-frequency words undergo more change or change at a faster rate than low-frequency words, (p. 20)

[S]ubstantial evidence has recently been reported showing that phonetic reduction occurs earlier and to a greater extent in high-frequency words than in low-frequency ones...[I]f we postulate that reduction occurs online as words are used, then words that are used more often are exposed to reduction processes more often and thus undergo change at a faster rate. (p. 37)

Although grammatical words maybe realized by single wordforms, the correlation between these units is often disrupted by phonological or syntactic processes. Cliticization creates sequences in which multiple grammatical words correspond to a single phonological word. In example (2a), the contracted phonological word *he’s* corresponds to the grammatical words *he* and *has*. Separable particles illustrate the converse case, in which a single grammatical word such as German *ablehnen* “reject” is realized by the wordforms *lehnen* and *ab* in (2b).

#### (2)

- a. *He’s* cut the flowers.
- b. Sie *lehnen* den Vorschlag *ab*.  
they “lean” the proposal “away”  
“They reject the proposal.”

### Cues and correlations

The divergence between grammatical and phonological word has led some scholars to question the status and even the usefulness of the notion “word” in the analysis of a grammatical system.<sup>1</sup> Although this question has attracted renewed attention in the recent literature, it articulates a longstanding criticism of traditional word-based approaches. Thus the difficulty of demarcating words is taken by Bloomfield (1914) as evidence for the primacy of the sentence.

it has long been recognized that the first and original datum of language is the sentence—that the individual word is the product of a theoretical reflection that ought not to be taken for granted. (p. 38)

One need not dispute the epistemological priority of utterances or downplay the discrepancies between phonological and grammatical words to realize that these discrepancies arise precisely because there *are* phonetic cues that, with varying degrees of reliability, mark word boundaries or otherwise guide the segmentation of utterances into words. The existence of mismatches should not obscure the fact that the two notions of “word” overlap, at least partially, in many languages, and that this overlap permits speakers to isolate grammatical words.

Although grammatical words may be imperfectly demarcated, subword units—including, significantly, roots—are rarely if ever cued at all by phonetic properties. There is no discrepancy between the “grammatical morpheme” and the “phonological morpheme” for the simple reason that there is no such thing as a “phonological morpheme.” Hence the objection that grammatical words are not reliably and invariantly cued in the speech stream provides no motivation for shifting the focus of morphological analysis onto units smaller than the word (such as morphemes), since these units require an even greater degree of abstraction from the speech signal. The observation that words are abstractions just falls under the broader generalization that *all* linguistic units smaller than utterances are abstracted from larger sequences of connected speech.

It is also possible that much of the debate about the alignment of grammatical and phonological words is fundamentally misconceived. Nearly all linguistic approaches to this issue frame the problem in terms of specifying reliable and cross-linguistically valid cues for demarcating grammatical words. Underlying these approaches is the assumption that there should be some set of necessary and sufficient conditions for defining words in a given language, or across languages. However, in parallel to theoretical and descriptive studies, there exists a large and diverse psycholinguistic and computational literature on word segmentation and recognition. This literature includes the work on identifying “uniqueness points” in Marslen-Wilson and Welsh (1978) and Marslen-Wilson and Tyler (1980), neural network-based predictive models (Elman 1990), and statistical models of word segmentation (Goldwater, Griffiths, & Johnson, 2009). The models of word segmentation developed in these studies are based on the observation that entropy (roughly, uncertainty about the segments that follow) declines as more of a word is processed, then peaks again at word boundaries. These sorts of prediction-based approaches suggest that the search for invariant cues may be futile and that the observable phonetic properties that have been assumed to be definitional are secondary cues for what is an essentially statistical notion.

Observations about predictability at word boundaries are consistent with two different kinds of assumptions about what constitutes a word: either a word is a unit that is statistically independent of other units, or it is a unit that helps to predict other units (but to a lesser degree than the beginning of a word predicts its end).

(Goldwater et al., 2009, p. 22)

From this perspective, the phonological word might correspond to phonetic sequences between entropy “peaks,” with phonetic cues representing one source of entropy reduction. The correlation between phonological and grammatical words could be similarly probabilistic rather than defined by discrete criteria.

Grammatical and phonological words represent two kinds of word tokens. In examples (1) and (2), there are four occurrences of the wordform *cut*, corresponding to two occurrences of the past participle *cut*, and one occurrence each of the preterite and infinitive. Given that grammatical words are defined as wordforms with fixed meanings, their identification depends on assumptions about the parts of speech and morphosyntactic features in a language. For example, a classification that recognizes distinct perfect and passive participles in English will treat the two occurrences of *cut* in (3) as distinct grammatical words. On an account that recognizes a unitary past participle in perfect and passive constructions, there will be two occurrences of a past participle *cut*. More generally, the identification of grammatical words is always defined relative to a grammatical analysis (or, in the case of corpora, a set of grammatical “tags”).

**(3)**

- a. He has *cut* the flowers. (PERFECT/PAST PARTICIPLE)
- b. The flowers have been *cut*. (PASSIVE/PAST PARTICIPLE)

### Lexemes and Lemmas

Insofar as occurrences of grammatical words are tokens of a common word “type,” they imply another, more abstract notion of “word.” This notion of “word type” usually goes under the name LEXEME. Matthews (1972, p. 160) characterizes the lexeme in this sense as “the lexical element...to which the forms in [a] paradigm as a whole...can be said to belong.”<sup>2</sup> In a later discussion of the same point, Matthews (1991, p. 26) suggests that a lexeme is “a lexical unit and is entered in dictionaries as the fundamental element in the lexicon of language.” Lexemes are thus reminiscent of the lexicographer’s notion of a LEMMA, which is the citation form of an item or the headword under

which it is listed in a dictionary.<sup>3</sup> The connection between these notions is reinforced by the fact that lexemes are conventionally represented by the citation form of an item in small caps (i.e., by the lemma of the item). However, whereas a lemma is a distinguished FORM, for example, the infinitive *cut*, a lexeme is a set of grammatical WORDS.

The contrast between wordforms, grammatical words, and lexemes resolves a systematic ambiguity in the use of the term “word” as applied to the pure form *cut*, the preterite *cut*, and the lexeme *CUT*. Not all items give rise to a full three-way split, since different notions of word may coincide in particular cases. There is usually at least a partial correlation between grammatical words and wordforms. For closed-class categories, the distinction between lexeme and grammatical word may not be especially relevant or useful, since a preposition or conjunction will usually be associated with a single grammatical word. Much the same will be true of open-class categories in an isolating language, such as Vietnamese, if nouns and verbs are represented by single grammatical WORDS.

### Paradigms and Families

Determining which grammatical words are part of a lexeme also introduces the kinds of issues that confront lexicographers when they attempt to distinguish primary “word entries” from “word senses” in a language. These issues bear in a direct way on the delineation of inflectional and derivational processes. From a traditional perspective, inflectional processes are said to define (or, more generally, relate) forms of a lexeme (i.e., grammatical words), whereas derivation derives (or relates) lexemes. It is typically assumed that the grammatical words that comprise a lexeme must belong to the same word class, preserve a core lexical semantics and argument structure, and even share a set of “intrinsic” features that are invariant for a given item. Inflectional processes are accordingly regarded as monotonic processes that add features to the invariant properties that can be associated with stems. In contrast, processes that alter class or argument structure are treated as derivational, since these changes create new lexemes.

Yet traditional descriptive practices are not entirely consistent with this split, as noted in a number of recent works, including Booij (1996), Haspelmath (1996), and Spencer (1999). In particular, conjugational paradigms often include categorial hybrids, such as participles, which exhibit adjectival properties, and gerunds or masdars, which exhibit nominal properties. The inclusion of passive participles also introduces argument structure variation into conjugational paradigms. This inconsistency undermines the coherence of the notion “lexeme,” along with associated conceptions of the inflection/derivation divide, if, as these authors assume, the notions “lexeme” and “inflectional paradigm” are taken to coincide.

Alternatively, one can interpret traditional descriptions as inexplicit rather than inconsistent, and reconstruct lexemes as sets of forms that occupy an intermediate position between inflectional paradigms and morphological families. This interpretation has the additional benefit of resolving a systematic ambiguity in the use of the term “paradigm.” Nouns and adjectives typically have a unified set of forms, inflected for case, number, definiteness, and so forth, which can be consolidated within a single declensional paradigm. In contrast, the inflected forms of verbs often exhibit a suborganization into sets of forms that share the same tense/aspect/mood features and vary in their agreement features. The term “paradigm” is applied to these smaller sets when one refers to the present indicative paradigm of French *PARLER* “to speak” and to the full set of inflected forms when one refers to the paradigm of *PARLER toute court*. To resolve this ambiguity, the term “paradigm” can be reserved for the first, smaller set of forms and the term “lexeme” applied to the complete set of inflected forms. Hence the present indicative paradigm of *PARLER* comprises a subset of the lexeme *PARLER*. The third member in this concentric classification is the “morphological family,” which contains a lexeme, such as *PARLER*, and any lexemes-related *PARLER*.

A coherent split between inflection and derivation can then be maintained by assigning related forms to the appropriate form class associated with an item. Consider first the challenges presented by categorial hybrids. If active participles are treated as neutralizing the categorial distinction between verbs and adjectives, they can be included in the lexeme of a verb provided that the lexeme is defined as a set of grammatical words with NONCONTRASTIVE (rather than identical) values for word class and other intrinsic features. A similar analysis extends to other elements that neutralize rather than alter word class distinctions. However, a derived deverbal nominal will, irrespective of how productively it is formed, comprise a separate lexeme in the morphological family of a verb.

The classification of passive forms likewise depends on the treatment of argument structure. Passive forms will be included in the lexeme of a verb if the forms in a lexeme need only share a common logical (or thematic) argument structure. If the words that make up a lexeme must also have common surface valence (subcategorization)

demands, then active and passive words will belong to different lexemes but again to a common morphological family. Either choice restores coherence to a traditional perspective in which lexemes occupy a position between inflectional paradigms and morphological families.<sup>4</sup> A lexeme such as *CUT* will contain a set of grammatical words, including finite and nonfinite verb forms. However, *CUT* can itself be located within a morphological family that includes the predictable agentive nominal lexeme *CUTTER*, which contains the singular grammatical word *cutter* and its plural counterpart *cutters*.

### Endocentric and Exocentric Processes

Given a systematic contrast between lexemes and morphological families, the somewhat overloaded contrast between “inflection” and “derivation” can be construed in terms of a distinction between *ENDOCENTRIC* or “within-lexeme” processes, and *EXOCENTRIC* or “cross-lexeme” processes. Characteristics that correlate with the endocentric-exocentric split need not then enter into the *DEFINITION* of this contrast. The fact that inflection operates over a closed, relatively uniform space makes inflectional processes interpredictable and productive. From the class of an open-class item, it is usually possible to determine the features that are distinctive for the class and predict the number of forms of the item; apart from irregular items, paradigms and lexemes are comparable in size and structure within a word or inflection class.<sup>5</sup> The closed, uniform feature space of an inflectional system thus defines a “grid” populated by interpredictable forms.<sup>6</sup>

The uniform organization of inflectional paradigms contrasts starkly with the more variable structure exhibited by families of derivational forms. Exocentric processes that can change word class, valence, or other intrinsic properties clearly do not define a finite set of forms within a uniform feature space. From the class of an item, one cannot in general predict the number and type of derivational formations in which it occurs. Given a list of derivational processes active in a language, it is of course possible to assign a uniform family of “potential” forms to all of the members of a word class. Yet the uniformity achieved is deceptive, because it collapses a critical distinction between those forms that are *ESTABLISHED* in a language and those that are merely possible in principle. The status of derivational forms that are possible but unattested is similar to that of acceptable phoneme sequences that “merely” lack a conventionalized meaning. Moreover, defining expanded derivational paradigms does not make it any easier for a speaker to predict the derivational formations that are attested and in use within a language. The issue is perhaps clearer in connection with compounds. Of the infinitely many possible noun compounds in English only a comparatively small number are established, and a speaker cannot predict the set of established compounds containing an item from the item itself.

These contrasts between inflection and derivation are reflected in the ways that morphological systems are described and analyzed. It is rare for an individual inflected form to be described as established on its own. The availability of a form correlates instead with the productivity of a whole pattern. Conversely, notions like “morphological gap,” “suppletion,” and even, to a large degree, “syncretism” are mainly or exclusively applied to inflectional paradigms. Derivational families do not usually have “gaps” because they do not define implicational relations over a closed and uniform space that give rise to expectations about the existence of specific forms, even where the shape would be predictable. Suppletion can likewise only arise where there are definite assumptions about the shape of particular members of a form set. Whereas inflectional paradigms generate strong assumptions of this nature, derivational families do not. Syncretism presupposes a similar structure, as syncretic patterns imply the existence of independent cells that can be associated with fully or partially identical forms. But the notion of a derivational cell is not clearly defined in the absence of features that specify a morphosyntactic grid within which to place the cells.

Although there is a strong correlation between endocentricity and predictability, it is neither possible (nor desirable) to define inflectional/endocentric processes in terms of predictability or productivity. It is possible for endocentric processes to be unpredictable and for exocentric processes to be predictable. Neither of these patterns raises practical or theoretical difficulties. An unpredictable inflected form is usually described as sporadic or lexically idiosyncratic and must, in the general case, be learned on an item-by-item basis. The “past subjunctive” forms of German fall into this category, as do comitative, abessive, or instructive case forms in Finnish. Exocentric processes can likewise be highly predictable while defining new lexemes, with their own inflectional paradigms.

For example, agentive nominals in *-er* can be formed for virtually any verb in English, so that the predictability of

agentive nominals in *-er* approaches that of inflected verb forms. Nevertheless, there are two important differences between an agentive nominal, such as *WALKER*, and an inflected form, such as *walks*. The first is that *WALKER* is associated with its own nominal paradigm, containing singular *walker* and plural *walkers*. Treating the process that defines *-er* nominals as inflectional/endocentric would make the lexeme *WALK* a categorially incoherent collection of inflected verb and noun forms. The second difference is that inflected forms are associated with a specific paradigm within a lexeme, whereas derivational forms tend to be associated with the lexeme. This distinction can be illustrated even in the simple conjugational system of English. Whereas *walks* is associated with the present indicative paradigm, *walker* is not the agentive nominal of any particular paradigm but of the verb lexeme *WALK*.

## Implications

The contrast between wordforms, grammatical words, and lexemes clarifies the principal senses of the notion “word.” The distinction between paradigms, lexemes, and families in turn defines the paradigmatic organization relevant to the split between endocentric and exocentric morphological processes. The resulting classification reconstructs a traditional perspective in a form that is compatible with, and even to some degree supported by, recent psycholinguistic studies.

Traditional claims about the “psychological reality” of words and paradigmatic structure receive a measure of support from experimental studies of frequency effects on lexical processing. A number of studies have shown that the frequency of inflected forms and the size of derivational families have a robust effect on lexical processing. One line of research has investigated the correlation between response latencies for inflected forms and the token frequency of the elements of their inflectional paradigms (Taft, 1979; Baayen, Lieber, & Schreuder, 1997; Hay, 2001). A second line of research has demonstrated that the processing of an item is facilitated by the semantically transparent items in its morphological family and inhibited by the semantically opaque items (Schreuder & Baayen, 1997; de Jong, 2002; Moscoso del Prado Martín, 2003). A third line of research attempts to provide a single information-theoretic measure that subsumes the token-frequency effects relevant to inflectional paradigms and the type-frequency effects relevant to morphological families (Kostić, Marković, & Baucal, 2003; Moscoso del Prado Martín, Kostić, & Baayen, 2004). More recent studies (Milin, Filipović Đurdjević, & Moscoso del Prado Martín, 2009; Milin, Kuperman, Kostić, & Baayen, 2009) have also shown that speaker responses are conditioned in part by the relation between relative frequencies. In particular, speakers appear to be sensitive to the relation between the probability distribution of the inflected forms of a lexical item and the probability distribution of exponents within the inflection class to which the item belongs. The more that these distributions diverge, the longer are the response latencies and the higher are the error rates in lexical decision tasks.

The traditional contrast between inflection and derivation may likewise contribute to an explanation for the observation that the processing of inflectional forms is more sensitive to token frequencies, whereas the processing of derivational formations is more sensitive to the type frequencies (Baayen et al., 1997). The relevant assumption is just that members of a given paradigm or inflection class have a comparable number of forms, whereas the size of derivational families may vary by orders of magnitude. Hence, if the type frequencies for inflected forms are relatively constant, only token frequencies will be distinctive.

Experimental methods and results are described greater detail in later sections. However, two points are worth mentioning here. The first is that all available research suggests that speakers are sensitive to frequency and distributional relations, information that is almost completely absent from theoretical models. The second is that the processing of a form appears to be influenced (whether facilitated or inhibited) by related forms, an effect that conflicts with the theoretical idealization that individual forms are defined in isolation. Morphological family effects seem to indicate that related forms are not only present as elements of a speaker’s mental lexicon but that they are in some sense “activated” in the processing of a given form. These effects appear particularly compatible with models that link families of wordforms into networks of elements with shared formal, grammatical, and/or semantic properties (Bybee, 1985, 2010).

## Word Structure

The models of morphology that grew out the classical Greco-Roman tradition recognized no grammatical unit intervening between sounds and words, and did not even see the need to isolate roots (Law, 1998, p. 112).

However, since at least the early Neogrammarian period (Schleicher, 1859) the study of morphology has been centrally concerned with the morphotactic organization of words. The recurrent structural patterns exhibited by words in turn defined the familiar models of morphological typology initially proposed in Sapir (1921). The following sections summarize this organization, from patterns through types.

## Patterns

Morphotactic patterns fall broadly into affixal (concatenative) and nonaffixal (nonconcatenative) classes. Affixal patterns are characterized by the presence (or absence) of a formative. Nonaffixal patterns may be distinguished by a stress pattern, a tonal melody, a length contrast, an accentual pattern, or another change in the properties of a segment or sequence of segments.<sup>7</sup>

## Affixation

The most common type of affixation is suffixation. Thus the pattern represented by the suffixation of *-s* in English *books* is more common than the prefixation of *re-* in *reread*, and both are more prevalent than other types of affixation. INFIXATION tends to arise as a prosodically conditioned variant of prefixation or suffixation. Hence, few infixes do not have prefixal or suffixal variants. In English, additional emphasis can be placed on expletives, such as *bloody*, by interposing them between two prosodic feet of the noun that they are modifying: *bloody unlikely~un-bloody-likely* (McMillan, 1980; McCarthy, 1982). In Tagalog, the “actor focus” marker *um* is prefixed to vowel-initial verb root *alís* in (4a), but infixes after the initial consonant of the consonant-initial root *bása* in (4b).

**(4)** Infixation in Tagalog (Aspillera 1981, p. 45f.)

- a. *alís* ~ **um***alís* “to leave”
- b. *bása* ~ **bu***mása* “to read”

CIRCUMFIXATION (or paraffixation) is the most sparsely attested type of affixation, and many cases invite a reanalysis in terms of successive affixation. In German, for example, the formation of the perfect participle is sometimes described in terms of circumfixation. A regular verb like SAGEN “say” has a basic stem *sag-* and a perfect participle *gesagt*. However, there is no clear motivation for a circumfix *ge-...-t*, given that the preterite series is based on a stem *sagt*, so that *gesagt* can be formed by prefixing *ge-* to this stem (Blevins, 2003). More plausible cases of circumfixation are illustrated by the Georgian superlative marker *u...esi* in (5a) and the Chuckchee negative marker *e... ke* in (5b).

**(5)** Circumfixation in Georgian and Chuckchee

- a. *q’ru* ‘deaf’ ~ **u-q’ru-esi** “the most deaf” (Tschenkéli, 1958, p. 225)
- b. *tejkev-ək* “fight” ~ **e-tejkev-ke-it-ək** “not to fight” (Comrie, 1981, p. 247)

REDUPLICATION is sometimes regarded as a type of parasitic affixation that copies all or part of a base form. The future in Tagalog is formed by reduplicating the first syllable of a root, for example, *a* in *al.ís* in (6a), but *bá* in *bd.sa* in (6b).

**(6)** Reduplication in Tagalog (Aspillera, 1981)

- a. *alís* ~ **aalís** “leave.FUT”
- b. *bása* ~ **babása** “read.FUT”

## Nonaffixal Patterns

Nonaffixal patterns are fairly heterogeneous, and correspond to any phonetic contrast that is stable and discriminable enough to be morphologized. SUBTRACTION (or truncation) involves the loss of a segment, usually under prosodic conditions. The formation of Hidatsa imperatives in (7a) illustrates a subtractive pattern cited by Harris (1942, p. 110). The Papago forms in (7b) provide another familiar case, in which perfectives lack the final *-m* of the imperfectives.

**(7)** Truncation in Hidatsa and Papago

- a. *cicic* “he jumped” ~ *cic* “jump!”
- ika.c* “he looked” ~ *ika* “look!”

- b. *him* ~ *hi* “walking,”  
*hihim* ~ *hihi* “walking.PL”

(Zepeda, 1983)

In patterns of ABLAUT (or vowel modification), a morphological contrast is signaled by a change in vowel quality. In German, present, preterite, and participial forms of strong verbs exhibit the residual ablaut patterns in (8a) and (8b). In the Lezgian examples in (8c) and (8d), the inessive form of a noun “is marked by lowering the final vowel of the oblique stem” (Haspelmath, 1993).

- (8) Ablaut in German and Lezgian
  - a. *singe* “sing.1SG.PRES” ~ *sang* “1SG.PRET”  
~ *gesungen* “PART.”
  - b. *nehme* “take. 1SG.  
PRES” ~ *nahm* “1SG.PRET” ~ *genommen* “PART.”
  - c. *čarxú* “rock” ~ *čarxá* “INES.SG”
  - d. *arčí* “paper” ~ *čarčé* “INES.SG”

The phenomenon of morphological GRADATION is illustrated by Estonian, which exhibits a three-way quantity distinction between Q1 (short), Q2 (long), and Q3 (overlong) syllables. In the morphological system, the contrast between Q2 and Q3 distinguishes weak and strong (quantitative) grade. In the class of nouns represented in (9), the genitive and partitive singular forms differ solely in grade: a weak genitive singular contrasts with a strong partitive singular.<sup>8</sup>

- (9) Grade in Estonian declensions (Mürk 1997; Blevins, 2008)
  - a. *luku* “lock. GEN.SG” (Q2) ~ ‘*lukku*’ ‘lock.  
PART.SG (Q3)
  - b. *kooli* “school.GEN.SG” (Q2) ~ ‘*kooli*’ “school.  
PART.SG” (Q3)

Although the contrasts in these examples cannot be keyed to segmental formatives, the elements of each of the patterns can be associated with a particular morphological function or meaning. Even this constancy is missing in EXCHANGE patterns, in which a morphological contrast is encoded by a symmetrical or “reversible” opposition. In Spanish, the contrast between indicative and subjunctive mood is marked by a process of *a~e* “vowel reversal” (Matthews, 1991, p. 198) illustrated in (10a) and (10b). The contrast between the Estonian partitive

Figure 9.1 Contrastive modulation in Dinka. Adapted from Andersen (2002). Singulars and plurals in (10c) and

Alternation	Absolutive	Oblique	
ú~û	ájít	ájít	‘chicken’
û~ú	árêw	árêw	‘tortoise’

(10d) is encoded by an *i ~ e* alternation.

- (10) Exchange patterns in Spanish and Estonian
  - a. *compra* “buy.3SG.IND” ~ *compre*  
“buy.3SG.SUB”
  - b. *come* “eat.3SG.IND” ~ *coma* “eat.  
3SG.SUB”
  - c. ‘*kooli*’ “school.PART.SG” ~ ‘*koole*’ “school.  
PART.PL”
  - d. ‘*lille*’ “flower. PART.SG” ~ *lilli*  
“flower.PART.PL.”

# The Morphology of Words

Dinka exhibits a similar pattern of “variable interpredictability,” in which suprasegmental tonal contrasts distinguish absolutive from oblique case.

The oblique is distinguished from the absolutive in virtually all monosyllabic nouns that have a short vowel...and in most disyllabic nouns with the prefix à- and a short root vowel...The rule for such nouns is that if the absolutive has a low root tone...then the oblique gets a falling root tone, and if the absolutive has a high or falling root tone...then the oblique gets a low root tone.

(Andersen, 2002, p. 9)

The near-minimal contrast between àjít ~ àjít and àrêw ~ àrêw in Figure 9.1 illustrates the alternation between low-falling and falling-low tonal melodies.

METHATHESIS represents an extreme version of a nonaffixal strategy in which a morphological contrast is expressed by a process that rearranges the relative order of segments or formatives. In Rotuman (Churchward, 1940; Blevins & Garrett 1999, p. 510) CV metathesis distinguishes the CV “complete phase” (corresponding roughly to a definite form) from the VC “incomplete phase.”

(11)

- a. *seséva* ~ *seséav* “erroneous”
- b. *tíko* ~ *tíok* “flesh”

## Morphological types

Although different patterns and strategies often predominate in different parts of a language, it is common to classify languages by positioning them relative to a number of ideal morphological types. At one extreme are languages of the ISOLATING type, in which each morpheme functions as a grammatical word. Modern English approaches this ideal, at least insofar as inflectional morphology is concerned, although Southeast Asian languages, such as Vietnamese, are usually regarded as paradigm examples of the isolating type. At the other extreme are AGGLUTINATIVE languages, in which words consist of a sequence of discrete formatives, each of which realizes a single property. Many inflectionally complex languages contain structures that conform to this ideal. For example, a Finnish noun, such as *taloissansa* “in their houses,” can be assigned the agglutinative analysis in (12a). However, in few languages is this transparency maintained through the entire morphology. In Finnish, agglutination breaks down already in the nominative plural *talot* “houses” in (12b), which lacks the distinct plural marker *-i-* and the discrete case marker found in other forms.

(12) Agglutination in Finnish (Karlsson, 1999)

- a. *talo-i-ssa-nsa* “HOUSE-PL-INES-3PL.POSS”
- b. *talo-t* “HOUSE-NOM.PL”

re:k (STEM)	s	is	ti: 2ND SING PERFECTIVE
----------------	---	----	-------------------------------

[Click to view larger](#)

Figure 9.2 Morphological analysis of Latin *re:ksisti:*. Adapted from Matthews (1972).

e	le	lý	k	e	te
		PERF		PAST IND	2ND PLU
Past		ACTIVE			

[Click to view larger](#)

Figure 9.3 Morphological analysis of Greek *elélékete*. Adapted from Matthews (1991).

The “fusion” of nominative case and plural number in the single suffix *-t* in (12a) is—like the realization of third person and plural by the ending *-nsa* in (12a)—a salient characteristic of languages of the (IN)FLECTIONAL type.

Fusional patterns are ubiquitous within Indo-European languages, and particularly well represented in Latin and Ancient Greek. As Matthews (1972, p. 132ff.) shows, the Latin verb *re:ksisti*: “you had ruled” (*rexisti* in the standard pedagogical orthography) exhibits the many-many relations between features and forms displayed in Figure 9.2. The ending *-ti*: exhibits a fusional pattern, realizing perfective aspect, second person and singular number. The perfective feature exhibits a converse “fis-sional” pattern, as it is realized by each of the formatives *-s-*, *-is-*, and *-ti*:

Nor is this an extreme or unrepresentative example. Classical Greek exhibits even more tangled feature-form associations, as illustrated by the verb *elelykete* “you had unfastened,” in Figure 9.3 below. Just as *re:ksisti*: occupies the second person perfective active cell of REGO “rule,” *elelykete* realizes the second person past perfective indicative active cell in the paradigm of LYO “unfasten.” But as Matthews (1972, 1991) observes, the scattered realization of aspect and voice confounds any attempt to impose an agglutinative structure on these forms. Simpler cases of extended exponence are illustrated by the combination of a suppletive future stem *ir-* and a future ending *-ai* in French *irai* “go.1SG.” More “exuberant” patterns are likewise attested in Batsbi, as shown by Harris (2009).<sup>9</sup>

## Morphological analysis

It may be useful to bring this summary of the linguistic properties of words to a close with a brief discussion of some of the analytical issues that arise in assigning wordforms a morphotactic structure or associating them with sets of paradigmatic alternatives. Interestingly, the interaction between paradigmatic alternatives and morphotactic structure plays a significant role in morphological systems. The implicit “competition” between alternatives gives rise to a type of “paradigmatic deduction” that allows forms to be interpreted based on the absence of a marker. Noun paradigms in English provide a simple example. Plural nouns in English are marked by the suffix /z/ (represented orthographically as -s). There is, however, no marking of singular number, and none is needed, given that a singular noun is unambiguously identified by the lack of a plural marker. A zero morph adds no information to what speakers can already deduce from the absence of any realized exponent. Similar patterns are even more typical of more intricate paradigms. As Anderson (1992, p. 87) notes, Georgian verbal paradigms provide a striking illustration of the fact that that “information may sometimes be conveyed not by constituents that are present in the structure of a given word, but precisely by the fact that certain other material is ABSENT.”

Cases of what are sometimes termed “gestalt exponence” likewise highlight the grammatical and semantic contribution of the morphotactic arrangement of formatives. The patterns of cumulative and extended exponence illustrated in the section on morphological types, together with the paradigmatic deduction described in English and Georgian, suggest that the properties of words are not, at least in any straightforward or direct way, reducible to the properties of their parts. A range of patterns cannot be made to conform to an agglutinative ideal in which the grammatical meaning of a word consists of discrete units of meaning that correspond to subword formatives. Instead, morphological systems tend to reuse common parts in different contexts in such a way that the arrangement of these parts contributes a “constructional” meaning. Expanding on the discussion of gradation in (9) will clarify how combinations of elements can have distinctive meanings within a language. The first four columns in Table 9.1 contain the singular grammatical case forms of nouns that exhibit productive “weakening” gradation in Estonian. The

Table 9.1. Singular nouns in Estonian. Adapted from Erelt, Erelt, and Ross (2000) and Blevins (2008). nominative

Nominative	Sukk	Kukk	Pukk	Lukk	Lugu
Genitive	Suka	Kuke	Puki	Luku	Loo
Partitive	Sukka	Kukke	Pukki	Lukku	Lugu
Illative 2	Sukka	Kukke	Pukki	Lukku	Lukku
	“Stocking”	“Rooster”	“Trestle”	“Lock”	“Tale”

forms of these nouns consist of a strong (Q3) stem, identified by the double consonant *-kk*. The remaining forms consist of a strong stem, or a weak stem in *-k*, followed by one of the “theme vowels” *a*, *e*, *i*, and *u*.<sup>10</sup>

Let us consider the locus of the property “partitive singular.” This property can only be associated with subword units if the forms in Table 9.1 are analyzable into smaller meaningful parts. The partitive singulars of this class contain two “recurrent partials”: a strong stem and a theme vowel. Thus *sukka* can be analyzed as *sukk* + *a*, *kukke* as *kukk* + *e*, *pukki* as *pukk* + *i*, and *lukku* as *lukk* + *u*. But partitive case cannot be associated either with strong stems or with theme vowels in isolation. The strong stems *sukk*, *kukk*, *pukk*, and *lukk* cannot be analyzed as partitive, because these same stems realize the nominative singular when they occur without a theme vowel, and also underlie the second “short” illative singular forms. Partitive case also cannot be associated with the theme vowels, because the same vowels occur in the genitive and illative singular forms.

Hence partitive case is an irreducibly word-level feature that is realized by the combination of a strong stem and a theme vowel. This type of gestalt exponence or “constructional” exponence (Booij, 2005) is difficult to describe if stems and theme vowels are represented separately. Because the grammatical meanings associated with strong stems and theme vowels are context-dependent, these elements cannot be assigned discrete meanings that “add up” to partitive singular when they are combined. From a traditional perspective, this context-dependence underscores the difference between “analyzability” and morphemic “decomposition.” An individual wordform is often analyzable into parts that recur elsewhere in its inflectional paradigm or in the morphological system at large. But these parts may function solely to differentiate larger forms, so that the minimal parts that distinguish a pair of wordforms cannot be associated with the difference in grammatical meaning between the wordforms. To return to the patterns in Table 9.1, the theme vowel *-u* distinguishes the partitive singular *lukku* “lock” from the nominative singular *lukk*. In isolation, however, the vowel *-u* neither realizes a specific case value nor expresses “the grammatical difference” between nominative and partitive. Exactly the same is true of the grade contrast between partitive singular *lukku* and genitive singular *luku*.

These examples illustrate some of the fundamental challenges that arise in decomposing wordforms into their component parts. Given a set of wordforms, it is often possible to segment them into roots or stems and derivational and inflectional exponents. Yet disassembling a word into its parts raises chronic and even recalcitrant problems in many languages. Perhaps the most extreme analytical problem is associated with the patterns of stem syncretism that Matthews (1972) calls “Priscianic” or “parasitic” and Aronoff (1994) later terms “morphomic.” A pair of Latin examples discussed in Matthews (1991) are summarized in the quotations below. The first case involves a correspondence between present active infinitives and imperfect subjunctives. The second example concerns the relation between past passive and future active participles.

For any Verb, however irregular it may be in other respects, the Present Infinitive always predicts the Imperfect Subjunctive. For the Verb “to flower,” *florere* → *florerem*; for the irregular Verb “to be,” *esse* → *essem*, and so forth without exception. (p. 195)

There are a few exceptions; but, in general, if the stem of the Past Participle is *x*, no matter how irregular it may be, that of the Future Participle is *x* with *-ūr-* added. (p. 200)

Table 9.2. Full paradigm of LUKK “lock.” Adapted from Erelt et al. (2000). The challenge posed by these patterns is

	Sing	Plu
Nominative	Lukk	Lukud
Genitive	Luku	Lukkude
Partitive	Lukku	Lukkusid
Illat2/Rart2	Lukku	Lukke
Illative	Lukusse	Lukkudesse
Inessive	Lukus	Lukkudes
Elative	Lukust	Lukkudest
Allative	Lukule	Lukkudele
Adessive	Lukul	Lukkudel
Ablative	Lukult	Lukkudelt
Translative	Lukuks	Lukkudeks
Terminative	Lukuni	Lukkudeni
Essive	Lukuna	Lukkudena
Abessive	Lukuta	Lukkudeta
Comitative	Lukuga	Lukkudega

that stems such as *florere* and *esse* (and past participle stems like *amāt* “to love”) are non-meaning-bearing units. Hence, once isolated, they cannot be assigned features and meanings that they contribute to determine the features and meanings of the larger units that they underlie. The difficulties that arise in attempting to assign these stems grammatical meanings merely highlights the fact that the value of these units of forms resides solely in the predictions that they sanction about other forms.

The descriptive value of form-based predictions is illustrated by the kind of implicational network that they define within the paradigm of LUKK “lock” in Table 9.2. The nominative singular *lukk* underlies the partitive singular *lukku*, which is in turn identical to the short illative singular *lukku* and underlies the partitive plural *lukkusid* and the genitive plural *lukkude*. The genitive singular *luku* underlies the nominative *lukud* and the illative through comitative singular, whereas the genitive plural *lukkude* underlies the illative through the comitative plural.

## Summary

In sum, the word occupies a central position in the morphotactic, paradigmatic, and implicational organization of a language. The traditional assignment of words to paradigms, lexemes, and families complements the analysis of wordforms into smaller parts, ultimately reflecting interrelated “outward-looking” and “inward-looking” perspectives on the most basic grammatical unit of a language.

## References

- Andersen, T. (2002). Case inflection and nominal head marking in Dinka. *Journal of African Languages and Linguistics*, 23, 1–30.
- Anderson, S. R. (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.
- Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.
- Aspillera, P. S. (1981). *Basic Tagalog*. Manila: M & L Licudine Enterprises.
- Baayen, R. H., Lieber, R., & Schreuder, R. (1997). The morphological complexity of simple nouns. *Linguistics*, 35, 861–877.
- Blevins, J., & Garrett, A. (1999). The origins of consonant-vowel metathesis. *Language*, 74, 508–556.
- Blevins, J. P. (2003). Stems and paradigms. *Language*, 79, 737–767.
- Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics*, 42, 531–573.
- Blevins, J. P. (2008). Declension classes in Estonian. *Linguistica Uralica*, 44, 241–267.
- Bloomfield, L. (1914). Sentence and word. *Transactions of the American Philological Society*, 45, 65–75. Reprinted in Hockett (1970), 38–46.
- Bloomfield, L. (1926). A set of postulates for the science of language. *Language*, 2, 153–164. Reprinted in M. Joos (1957), *Readings in linguistics I*. Chicago: University of Chicago Press (26–37).
- Bloomfield, L. (1933). *Language*. Chicago: University of Chicago Press.
- Booij, G. (1996). Inherent versus contextual inflection and the split morphology hypothesis. In G. Booij & J. van Marie (Eds.), *Yearbook of morphology 1995* (pp. 1–16). Dordrecht: Kluwer.
- Booij, G. (2005). Compounding and derivation: Evidence for constructional morphology. In W. U. Dressier, P. Kastovsky, & F. Rainer (Eds.), *Morphology and its demarcations* (pp. 109–132). Amsterdam: John Benjamins.
- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- Bybee, J. L. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, J. L. (2010). *Language, usage and complexity*. Cambridge: Cambridge University Press.
- Carstairs-McCarthy, A. (1994). Inflection classes, gender, and the principle of contrast. *Language*, 70, 737–788.
- Carstairs-McCarthy, A. (2005). Affixes, stems and allomorphic conditioning in Paradigm Function Morphology. *Yearbook of Morphology*, 2, 253–281.
- Churchward, C. M. (1940). *Rotuman grammar and dictionary*. Sydney, Australia: Methodist Church of Australasia.
- Comrie, B. (1981). *The languages of the Soviet Union*. Cambridge: Cambridge University Press.
- de Jong, N. H. (2002). *Morphological families in the mental lexicon* (Unpublished doctoral thesis). University of Nijmegen, Nijmegen, the Netherlands.
- Dixon, R. M. W., & Aikhenvald, A. Y. (Eds.). (2002). *Word: A cross-linguistic typology*. Cambridge: Cambridge University Press.
- Dromi, E. (1987). *Early lexical development*. Cambridge: Cambridge University Press.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.
- Erelt, M., Erelt, T., & Ross, K. (2000). *Eesti keele käsiraamat*. Tallinn: Eesti Keele Sihtasutus.

# The Morphology of Words

---

- Goldwater, S., Griffiths, T. L., & Johnson, M. (2009). A Bayesian framework for word segmentation: Exploring the effects of context. *Cognition*, 112, 21–54.
- Harris, A. C. (2009). Exuberant exponence in Batsbi. *Natural Language and Linguistic Theory*, 27, 267–303.
- Harris, Z. S. (1942). Morpheme alternants in linguistic analysis. *Language*, 18, 169–180. Reprinted in M. Joos (1957), *Readings in linguistics I*. Chicago: University of Chicago Press (109–115).
- Haspelmath, M. (1993). *A grammar of Lezgian*. Berlin: Mouton.
- Haspelmath, M. (1996). Word-class-changing inflection and morphological theory. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1995* (pp. 43–66). Kluwer, Dordrecht.
- Hay, J. (2001). Lexical frequency in morphology: Is everything relative? *Linguistics*, 39, 1041–1070.
- Hockett, C. F. (1958). *A course in modern linguistics*. New York: MacMillan.
- Hockett, C. F. (Ed.). (1970). *A Leonard Bloomfield anthology*. Chicago: University of Chicago Press.
- Karlsson, F. (1999). *Finnish: An essential grammar*. London: Routledge.
- Kostić, A., Marković, T., & Baucal, A. (2003). Inflectional morphology and word meaning: Orthogonal or co-implicative domains? In R. H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 1–44). Berlin: Mouton de Gruyter.
- Law, V. (1998). The Technē and grammar in the Roman world. In V. Law & I. Sluiter (Eds.), *Dionysius Thrax and the Technē Grammatikē* (pp. 111–120). Münster: Nodus Publikationen.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1–71.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29–63.
- Matthews, P. H. (1972). *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation*. Cambridge: Cambridge University Press.
- Matthews, P. H. (1991). *Morphology*. Cambridge: Cambridge University Press.
- McCarthy, J. J. (1982). Prosodic structure and expletive infixation. *Language*, 58, 574–590.
- McMillan, J. (1980). Infixing and interposing in English. *American Speech*, 55, 168–177.
- Milin, P., Filipović Đurdjević, D., & Moscoso del Prado Martín, F. (2009a). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language*, 60, 50–64.
- Milin, P., Kuperman, V., Kostić, A., & Baayen, R. H. (2009b). Words and paradigms bit by bit: An information-theoretic approach to the processing of inflection and derivation. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar: Form and acquisition* (pp. 214–253). Oxford: Oxford University Press.
- Moscoso del Prado Martín, F. (2003). *Paradigmatic structures in morphological processing: computational and cross-linguistic studies* (Unpublished doctoral thesis). University of Nijmegen, Nijmegen, the Netherlands.
- Moscoso del Prado Martín, F., Kostić, A., & Baayen, R. H. (2004). Putting the bits together: An information-theoretical perspective on morphological processing. *Cognition*, 94, 1–18.
- Mürk, H. W. (1997). *A handbook of Estonian: Nouns, adjectives and verbs*. Indiana University Uralic and Altaic Series, v. 163. Bloomington, IN: Indiana University.

# The Morphology of Words

---

- Nida, E. A. (1949). *Morphology: The descriptive analysis of words*. Ann Arbor, MI: University of Michigan Press.
- Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, 34, 191–243.
- Paul, H. (1920). *Prinzipien der Sprachgeschichte*. Tübingen: Max Niemayer Verlag.
- Robins, R. H. (1959). In defence of WP. *Transactions of the Philological Society*, 58, 116–144. Reprinted in *Transactions of the Philological Society*, 99, 116–144.
- Sadler, L., & Spencer, A. (1998). Morphology and argument structure. In A. Spencer & A. M. Zwicky (Eds.), *Handbook of morphology* (pp. 206–236). Blackwell.
- Sapir, E. (1921). *Language*. New York: Harcourt Brace.
- Schleicher, A. (1859). *Zur Morphologie der Sprache. Volume 1, Memoires de L'Académie de St Pétersbourg*. Leipzig: Leopold Voss.
- Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, 37, 118–139.
- Spencer, A. (1999). Transpositions and argument structure. In G. Booij & J. van Marie (Eds.), *Yearbook of morphology 1998* (pp. 73–101). Dordrecht: Kluwer.
- Stump, G. T. (2001). *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.
- Stump, G. T. (2005). Some criticisms of Carstairs-McCarthy's conclusions. *Yearbook of Morphology*, 2, 283–303.
- Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, 7, 263–272.
- Tschenkéli, K. (1958). *Einführung in die georgische Sprache*. Zurich: Amirani Verlag.
- Zepeda, O. (1983). *A Papago grammar*. Tucson, AZ: University of Arizona Press.

## Notes:

- (1) . See the papers in Dixon and Aikhenvald (2002) for extended discussion of this issue.
- (2) . An earlier use of “lexeme” is found in Hockett (1958, p. 169ff.), who uses the term to designate sequences that always occur as grammatical forms in a context where they are not part of any larger unit that also invariably occurs as a grammatical form. This usage is now largely obsolete.
- (3) . There is an alternative interpretation of the term “lemma” still in circulation. In psycholinguistic models of speech production, lemmas are often construed as abstract conceptual entries that represent “the nonphonological part of an item's lexical information” (Levelt, 1989).
- (4) . Moreover, on either alternative, “morphosemantic” alternations (Sadler & Spencer, 1998) like causativization will relate distinct lexemes within a larger morphological family.
- (5) . This predictability underlies the descriptive success of the realization-based models of inflection set out in Matthews (1972, 1991), Anderson (1992), Aronoff (1994), and Stump (2001).
- (6) . The patterns exhibited by this network provide the analogical base for exemplar-based models of both the traditional (Paul, 1920) and modern (Bybee, 2010) variety.
- (7) . Some theories of morphology regard affixal and nonaffixal patterns as different in character (Carstairs-McCarthy, 1994, 2005). Other approaches argue that “there is no theoretically significant difference between concatenative and non-concatenative inflection” (Stump, 2005, p. 284).

(8) . Since the Q3 quantity of partitive singular *kooli* is not marked orthographically (as by the consonant doubling in *lukku*), Q3 is marked diacritically by a single quotation mark.

(9) . The POLYSYNTHETIC patterns found in many Amerindian languages represent an orthogonal dimension of morphological variation, as these patterns essentially exhibit a distinctive division of grammatical labor. In many of the familiar languages of Europe and Asia, lexical meaning is expressed by roots or stems, and grammatical meanings by affixes and other types of exponents.

(10) . The choice of theme vowel is a lexical property of an Estonian noun, and is not conditioned by phonological properties of the stem or by morphosyntactic features of the noun.

### James P. Blevins

James P. Blevins is Assistant Director of Research at the Research Centre for English and Applied Linguistics in the University of Cambridge. He received his PhD in Linguistics from the University of Massachusetts, Amherst in 1990 and has published on a range of grammatical topics. His research deals mainly with the analysis of morphological systems and syntactic constructions, with an emphasis on paradigm structure and discontinuous dependencies. Areal interests include Germanic, Balto-Finnic, Balto-Slavic, and Kartvelian.

