CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

1 *Canonical Morphotactics*

1.1 Introduction and Objectives

A language's morphology defines two kinds of things. At the level of pure content, it specifies the grammatical categories and properties that its words embody (such as 'plural number', 'agent nominalization', and 'first conjugation') and the combinations into which these categories and properties enter (such as the combination '3rd person plural imperfect indicative passive, first conjugation' embodied by the Latin verb form *laudābantur* 'they were praised'). At the complementary level of morphological realization, a language's morphology identifies the formal exponents of its grammatical categories and properties (affixes, patterns of ablaut, grammatical tone or accent specifications, and so on) and the ways in which these exponents combine with a word's stem in realizing a word form's lexical and grammatical content. Intuitively, a language's ***morphotactics*** is that part of its morphology that determines the morphological realization of word forms – the system of principles that defines the patterns according to which word forms' grammatically significant parts are arranged as well as the relations among such patterns.[1]

Simple word forms naturally have simple morphotactics: the word form *dogs*, for example, is just the noun stem *dog* followed by the plural suffix *-s*. In such cases, it is hard to argue that the full word form isn't simply listed lexically (or, in psycholinguistic terms, stored and accessed as a unit), with its stem+affix structure simply built in. But as one examines word forms of increasing complexity, it becomes less and less reasonable to assume that a language's word forms are in all cases listed as wholes. A language's morphotactics must comprise means of inferring complex word forms that aren't necessarily listed. Logically, there are many possible frameworks within which

[1] It should be noted at the outset that the concept of the morpheme is not a necessary part of this conception of morphotactics, a point to which I return below.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

such inferences might be organized, and choosing among these possibilities is essential for the formulation of an adequate theory of grammar.

One possibility is to assume that a language's morphotactics is based on a set of skeletal word structures into which stems and affixes can be systematically inserted, either from the lexicon or by rules that associate stems and affixes with specific skeletal slots. (Such skeletal structures might be argued to be configurations of terminal nodes defined by syntax, or they might be defined by means of an autonomous morphological template.) On a word-skeletal approach of this sort, a language's morphotactic patterns are based on two different kinds of elements: concrete units with phonological content and abstract skeletal structures that determine the kinds of combinations into which those units enter as an effect of their insertion (whether from the lexicon or by rule).

Another possibility is to assume that a language's morphotactics comprises rules of exponence (rules that realize specific grammatical content by means of specific formal exponents) and that the order of these rules' application alone suffices to determine a word form's morphotactics without the help of any predefined skeletal structure. In an exponence-driven theory of this sort, a morphologically complex word's form is purely and simply an expression of the rules that realize its grammatical content with exponents. In fact, two sorts of exponence-driven theories might be distinguished. On the one hand, one might assume that all rules of exponence are minimal, in the sense that each rule introduces a single exponent. On this view, the Latin verb stem *laudā-* 'praise' is related to the imperfect passive verb form *laudābantur* 'they were praised' through the ordered application of three minimal rules of exponence, the first introducing the imperfect suffix *-bā* (short allomorph *-ba*), a second introducing the third-person plural suffix *-nt*, and a third introducing the passive suffix *-ur*. In this approach (the ordered rule approach), a language's morphotactic patterns are again based on two kinds of elements: individual rules of exponence and an overarching specification of the order in which these rules apply in the definition of a full word form.

On the other hand, one might assume that rules of exponence may themselves combine to form more complex rules of exponence, so that the stem *laudā-* is related to *laudābantur* through the application of a single, complex rule that affixes *-bantur*, a rule that is itself the combination of three simpler rules. In this approach (the rule-combining approach), a language's morphotactics simply specifies its inventory of (simple or complex) rules of exponence. (Such specifications might, of course, involve a basic set of simple rules and a set of principles for combining simpler rules to form more complex rules.)

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

On first consideration, one might suppose that the word-skeletal approach, the ordered rule approach, and the rule-combining approach are all essentially equivalent. Yet, a variety of phenomena can be shown to favor the third, rule-combining approach. In this book, I identify several phenomena of this sort, and I show how they can be accounted for in the context of a new theory of morphotactics. The phenomena at issue include periodic multiple exponence, asymmetrical affix dependencies, exceptional affix order, apparently nonlocal conditions on affix distribution, the paradigmatic opposition of one affix to a sequence of affixes, affix polyfunctionality, affix sequences expressing holistic content, affix counterposition (prefixation to a suffix or suffixation to a prefix), and affix potentiation and counterpotentiation, including parasynthetic derivation.

My objectives in this book are therefore twofold. On the empirical side, I document the wide variety of complexities that arise in the analysis of a language's morphotactics. I propose ten characteristics that canonical morphotactics might be assumed to possess in unremarkable cases, then demonstrate the numerous ways in which morphological systems deviate from these characteristics. Some such deviations are genuine; other phenomena are merely apparent deviations, ultimately conforming to canonical criteria if certain well-motivated assumptions are made.

On the theoretical side, I motivate an exponence-driven theory of morphotactics. I distinguish exponence-driven approaches to modeling a language's morphotactic characteristics from the word-skeletal approach, according to which a word form embodies an abstract skeletal structure or template that logically predefines the order of the particular exponents that give that word form its phonological substance. On the word-skeletal approach, the definition of a language's morphotactics is, in essence, the definition of its word forms' skeleta; many ways of defining such skeleta have been proposed (Selkirk 1982; Lieber 1992; Halle & Marantz 1993; Embick & Noyer 2001; Arregi & Nevins 2012; Crysmann & Bonami 2016). As I shall argue here, there are compelling reasons for favoring an exponence-driven approach to morphotactics over the word-skeletal approach.

In particular, I argue for a rule-combining approach to exponence-driven morphotactics, according to which a word's morphological form is the manifestation of an organized combination of rules of exponence. Crucially, I argue that the combinations into which rules of exponence enter are binary and potentially nested, and may involve any of four different modes of combination (namely composition, holistic combination, aggregation, and counterpotentiation).

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

I emphasize that the rules at issue here are not input–output rules for converting an underlying phonological representation to its superficial form, nor are they rules for constructing isomorphic representations of a complex word's form and content by putting morphemes together. Rather, the rules at issue here declare the organized expression of details of a complex word's content by details of its phonological form.

In the rule-combining approach to exponence-driven morphotactics, the central question is: What are the ways in which rules of exponence may combine in order to define the systematic aspects of the association of a complex word's content with its form? The null hypothesis is that in combining to define the association of a complex word's content with its form, rules of exponence always combine in the same way. I contend that this is not the case – that rules of exponence may in fact combine in at least four different ways in defining the correspondence between words' content and their form. Rules, in other words, are of different kinds. Simpler rules may combine to form more complex rules, and the patterns of such combination are various. Morphotactics, on this view, might instead be called ***regulatactics*** (< Latin *rēgula* 'rule'), the patterns of contact and arrangement among a word form's defining rules of exponence.

The rule-combining approach to exponence-driven morphotactics has important implications for both inflectional and derivational morphology. In the domain of inflection, the rule-combining approach makes it possible to regard every inflected word form as the expression of a single rule; in very many cases, this is a complex combination of simpler rules of inflectional exponence. I refer to a complex rule of this sort as a ***full exponence rule***; as I shall show, the definition of a language's system of full exponence rules may involve a variety of modes of rule combination, with ordinary rule composition being the default mode but by no means the only mode.

In the domain of derivation, the rule-combining approach makes it possible to postulate complex rules of derivation that result from the combination of simpler rules of derivational exponence. The existence of such complex rules is motivated by a variety of considerations. These include the phenomenon of potentiation; a contrary phenomenon of counterpotentiation; and instances in which a combination of derivational rules expresses more than the summed content of its component rules. As with inflection, the default mode of rule combination is that of composition, but the evidence discussed here motivates the postulation of additional modes.

The details of exponence-driven morphotactics are sometimes intricate, but the ultimate argument that motivates this approach is very simple: that a theory

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

of morphotactics requires rules of exponence, and by Occam's Razor, no further stipulations about the purported skeletal or templatic structure of words are needed, since rules of exponence in various combinations alone suffice to define a language's morphotactic patterns.

In this chapter, I discuss the preliminary assumptions of the rule-combining approach to morphotactics and advance the two fundamental hypotheses that underlie it: the morphotactic holism hypothesis and the morphotactic variety hypothesis (Section 1.2). In Section 1.3, I review previous proposals that provide empirical support for the morphotactic holism hypothesis, which (unlike the morphotactic variety hypothesis) is not a novel idea. In Section 1.4, I discuss the nature of canonical morphotactics, for which I introduce ten criterial characteristics, construed in rule-based terms. In Section 1.5, I give examples of phenomena that possess these characteristics as well as of phenomena that do not apparently possess them. The morphotactic phenomena to be analyzed in the following chapters deviate from some of these canonical characteristics but reinforce conformity to others provided that a rule-combining approach is assumed. In Section 1.6, I anticipate the range of topics to be discussed in subsequent chapters.

1.2 Rule-combining Morphotactics

The systematic word-internal relations exhibited by the components of a language's complex word forms constitute that language's morphotactics. In the morpheme-based approaches to morphology popular in the mid-twentieth century, a language's morphotactic principles are seen as constraints on the concatenation of a word form's morphemes. By contrast, rule-based conceptions of morphology represent a language's morphotactic principles as constraints on the interaction of the rules of exponence by means of which a word's form is defined. In this book, I examine a wide range of morphotactic phenomena in a variety of languages. The systematic patterns embodied by this evidence necessitate a more richly structured conception of the nature of morphotactics than follows from current assumptions (whether these be articulated in morpheme-based or rule-based terms). I therefore propose a new set of assumptions about morphotactics that is motivated by the need to provide an explanatory account of these numerous phenomena.

I develop this new conception of morphotactics in the context of a rule-based perspective on morphology. I favor the rule-based approach for reasons of descriptive precision, but the exponence-driven conception of morphotactics developed here could be profitably adapted to certain other

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

kinds of formal frameworks; for example, it has clear implications for the refinement of a construction-based approach to morphology (a point to which I shall return).

The rules at issue throughout the following discussion are declarative in nature – that is, they are clauses in the static definition of a language's morphology. They are not rules for converting underlying forms into superficial forms, nor are they rules for combining smaller form/meaning pairings to build larger form/meaning pairings. Rather, they are systematic generalizations of two kinds.

In the domain of inflection, a rule of exponence is a generalization about the relation between a word form's abstract representation and its concrete representation. Consider, for example, the two representations of the word *shoes* in (1). Elsewhere (Stump 2016), I have referred to an abstract representation such as (1a) as a content cell: a word form's content cell is the pairing of the lexeme that it expresses (e.g. SHOE)[2] with the morphosyntactic property set that it expresses (e.g. {plural}); a word form's content cell is what determines its syntactic distribution and its semantic interpretation. Lexemes are not units of form. Rather, they are units that possess lexicosemantic properties and are realized by stems and word forms; the lexeme SHOE, for example, possesses the lexicosemantic property 'count noun', denotes the set of shoes, and is realized by the stem *shoe-* as well as by the word forms *shoe* and *shoes*. The relation between the abstract representation in (1a) and its concrete representation (1b) is expressed by a default rule of exponence according to which ⟨L, {plural}⟩ is realized as ⟨Z-*s*, {plural}⟩, where Z is L's stem.

(1) Two representations of the word *shoes*
   a. Abstract: ⟨SHOE, {plural}⟩
   b. Concrete: ⟨*shoes*, {plural}⟩

In the domain of derivation, a rule of exponence is a generalization about the relation between the abstract and concrete characteristics of one lexeme and those of a related lexeme. Consider, for example, the lexemes SHOE and SHOELESS: the former has the property 'count noun', denotes the set of shoes, and has *shoe-* as its stem; the latter has the property 'privative adjective', denotes the set of things lacking shoes, and has *shoeless* as its stem. The relation between SHOE and SHOELESS is expressed by a default rule of exponence according to which a count noun L_1 with stem Z is related to a privative adjective L_2 with stem Z-*less*.

[2] Throughout, a lexeme's name is given in small caps.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

In formal terms, rules of exponence should be likened to mathematical functions: just as the squaring function f^2 denotes the relation between 2 and 4, between 3 and 9, between 3.5 and 12.25 (and so on), so the rule ⟦-**ed**⟧ in (2) denotes the relation between the weak stem /pɪk/ and the past-tense form /pɪkt/, between the weak stem /hʌm/ and the past-tense form /hʌmd/, between the weak stem /æd/ and the past-tense form /ædəd/, and so on. As this example shows, generalizations about superficial forms may perfectly well incorporate abstractions such as 'weak' or 'voiceless'; doing so in no way diminishes their fundamentally declarative nature. (Throughout, I use ⟦**x**⟧ as the label for a rule that introduces the exponent *x*. I use this sort of label to represent both inflectional rules, according to which *x* is the exponent of some morphosyntactic property set, and derivational rules, according to which *x* is the exponent of some derivational category. I introduce formal representations for the definition of rules of inflectional and derivational exponence in Section 2.1.)

(2) ⟦-**ed**⟧: In the inflection of a weak verb stem /X/, the property set {past} is expressed as
/X/-**ed** (where /YZ/-**ed** = /YZəd/ if Z is an oral alveolar stop,
otherwise = /YZt/ if Z is voiceless,
otherwise = /YZd/).

Rules serve two functions in a language's morphology.[3] First, they define form/content relations that are not listed in the lexicon. For example, an inflectional rule may define the form/content pairings of completely regular forms that are not frequent enough in their use to have induced lexical listing (e.g. forms such as the plural *shrikes* or the present participle *repaving*); in the derivational domain, a rule might be employed to create a novel lexeme such as BAGELIZE, which I have just now made up for use in sentences such as *They bagelized the traditional grilled-cheese sandwich*.[4]

Not all rules of exponence are sufficiently productive to be used to define form/content relations that are not listed in the lexicon. But even those that are not may still serve to simplify the lexicon. As Bochner (1993) has argued, lexical items that conform to rules are less "costly" to learn and to store than lexical items that deviate from rules in one way or another. For instance, verbs such as *ripen*, *sweeten*, and *toughen* conform to a regular (if unproductive) pattern that allows their lexical entries to be simplified in relation to those of the

[3] Jackendoff & Audring (2020: 52) make fundamentally this same point in distinguishing between the generative and relational uses of a morphological schema.

[4] Google shows that I am not the first person to have used the ⟦-**ize**⟧ rule to create a novel lexeme BAGELIZE.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

adjectives *ripe*, *sweet*, and *tough*; by contrast, the verb *christen* conforms less well to this pattern, so that its lexical entry requires more special stipulations.[5]

I argue here that a satisfactory synchronic account of a language's morphotactics must be one that allows the (potentially recursive) combination of simpler rules of exponence to form more complex rules of exponence. In my discussion, I pursue two fundamental hypotheses. The first of these, the ***morphotactic holism hypothesis*** in (3), is not a new idea, as I shall show presently. The second of them, the ***morphotactic variety hypothesis*** in (4), is an idea that has never before been investigated in a systematic way.

(3) Morphotactic holism hypothesis
A combination of rules of exponence may possess characteristics that do not follow from the characteristics of its simpler component rules.

(4) Morphotactic variety hypothesis
Rules of exponence may enter into different kinds of combinations possessing different characteristics.

In Section 1.3, I review a number of past proposals that support the morphotactic holism hypothesis. In Sections 1.4–1.5, I show that the content of hypotheses (3) and (4) can be most clearly elucidated in the context of a set of assumptions about the canonical characteristics of a language's morphotactics. As I shall show, such canonical characteristics are of two main kinds: those from which the postulation of morphological rule combinations constitutes a noncanonical deviation, and those that are reconciled with apparent deviations through the postulation of morphological rule combinations.

1.3 Past Work Supporting the Morphotactic Holism Hypothesis

A number of people have, in past work, suggested that the combination of two rules (or of two affixes) sometimes possesses characteristics that are not predictable from the characteristics of the individual rules (or affixes) constituting that combination.

Studies of the morphology of the Romance languages have long drawn attention to the phenomenon of parasynthesis (Darmesteter 1874: 80ff.; Corbin 1980; Scalise 1986: 21, 147ff.; Fradin 2003: 288–292; Fábregas & Scalise 2012: 62f.; Serrano-Dolader 2015). The label 'parasynthesis' is applied to two distinct but related phenomena. On the one hand, parasynthetic compounding

[5] Derivational rules may also simplify the lexicon by serving a stem-defining function, applying to a lexeme's "substem" to yield that lexeme's stem (see Section 2.1).

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

(Bisetto & Melloni 2007) refers to patterns of compounding that involve the integral presence of an affix (e.g. the suffix *-olo* in Italian *pesci-vend-olo* 'fish seller'). On the other hand, parasynthesis may also refer to derivational patterns such as that of Italian *imburrare* 'to butter', whose derivation involves conversion of the stem of the noun *burro* 'butter' to a verb stem (cf. *zuccher-o* 'sugar' → *zuccher-are* 'to sugar, sweeten'), but involves the integral presence of the prefix *in-* (Masini & Iacobini 2018). What these two kinds of patterns have in common is that they involve the inseparability of two logically distinct morphological operations. In *pescivendolo*, the compounding of *pesci* 'fish' goes hand-in-hand with the addition of the nominalizing suffix *-olo*; it is not plausible to think of *pescivendolo* either as the result of compounding *pesci* with a nonexistent noun **vendolo* (purportedly meaning 'seller') or as the result of suffixing *-olo* to the stem of a nonexistent verb **pescivendere* (purportedly 'to sell fish'). Similarly, it is not plausible to think of *imburrare* either as arising through the prefixation of *in-* to a nonexistent verb **burrare* or as arising by conversion from a nonexistent prefixed noun **imburro*; in *imburrare*, prefixation and conversion go hand in hand.

Bauer (1988) draws attention to the frequent incidence of cases in which two morphological markings work together to express a single piece of content. In English, for example, causative verbs are derived from adjectives in some cases by means of the suffix *-en* (e.g. *weak* → *weaken*), in other cases by means of the prefix *en-* (*able* → *enable*), and in still other cases by means of *-en* together with *en-* (*bold* → *embolden*). As Bauer shows, cooperative combinations of this kind not only involve all kinds of affixation but may also involve nonaffixal morphology; for example, the derivation of *bathe* from *bath* (/bæθ/ → /beɪð/) involves both ablaut (cf. *food* → *feed*) and consonant gradation (cf. *wreath* /riθ/ → *wreathe* /rið/). Bauer proposes the term 'synaffix' as a label for combined morphology of this sort; circumfixation and parasynthesis may accordingly be thought of as kinds of synaffixation. As Bauer shows, the components of a synaffix may or may not have uses apart from one another. Even if they do, labeling their combination a synaffix is nevertheless apt if the content expressed by their combination is not simply the composition of the content that they express when used separately; for instance, the Dutch synaffix *ge-* + *-te* expresses a collective meaning (*been* 'bone' → *ge-been-te* 'skeleton'), and although *ge-* and *-te* are used separately in other contexts, the collective meaning associated with the synaffix cannot be attributed to either affix on its own. In rule-based terms, this means that a combination of two rules may express holistic content that is not directly deducible from the rules' individual content.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-16821-2 — Morphotactics
Gregory Stump
Excerpt
More Information

Bochner (1993) argues that morphological simplicity is best assessed not by reference to a symbol-counting evaluation metric (which favors generalizations that minimize the amount of information that must be stored in the lexicon), but rather by reference to a pattern-matching metric (which favors generalizations that minimize the unpredictability of the information that is stored in the lexicon). In this context, he draws attention to many cases in which greater predictiveness can be attained by simultaneous reference to the patterns associated with distinct rules. Consider, for example, the rules in (5): (5a) defines deverbal adjectives in *-able* and (5b) defines deadjectival nominalizations in *-ity*. The fact that (5b) can in general apply to adjectives defined by (5a) might be expressed by means of the more specific rule in (5c), which makes simultaneous reference to the patterns associated with (5a) and (5b); (5c) is what Bochner terms a subset rule, since it specifies a subset of instances of the application of rule (5b). Rules (5a) and (5b) might be thought to make (5c) redundant in a morphological description of English; but in fact, (5b) is much less regular than the subset rule (5c). Rule (5b) isn't very regular; the adjectives *main* and *plain*, for example, are not nominalized as **manity* and **planity*. For this reason, the existence of *vanity* alongside *vain* in the English lexicon is not too predictable; by contrast, given the high regularity of (5c), the existence of *acceptability* alongside *acceptable* is highly predictable, a fact that presumably makes *acceptability* easier to learn and less "expensive" to store than *vanity*. Thus, as a kind of rule combination, (5c) contributes to the simplicity (i.e. to the predictability) of the English lexicon to an extent that (5a) and (5b) on their own do not.[6]

(5)

a. $\begin{bmatrix} /X/ \\ V \\ Z \end{bmatrix} \leftrightarrow \begin{bmatrix} /Xable/ \\ A \\ ABLE\ to\ be\ Zed \end{bmatrix}$

b. $\begin{bmatrix} /X/ \\ A \\ Z \end{bmatrix} \leftrightarrow \begin{bmatrix} /Xity/ \\ N \\ STATE\ of\ being\ Z \end{bmatrix}$

c. $\begin{bmatrix} /Xable/ \\ A \\ ABLE\ to\ be\ Zed \end{bmatrix} \leftrightarrow \begin{bmatrix} /Xability/ \\ N \\ STATE\ of\ being\ ABLE\ to\ be\ Zed \end{bmatrix}$

(Bochner 1993: 72, 88)

Stump (1993, 2001: 139–144) proposes that a portmanteau rule expressing a combination of inflectional categories α + β may compete with a

[6] In the context of somewhat different theoretical assumptions, Raffelsiefen 1992 and Booij 2010: 47–50 both argue for a similar conclusion.

© in this web service Cambridge University Press & Assessment www.cambridge.org