TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science (2018) 1–27

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12334

This article is part of the topic "2017 Rumelhart Prize Issue Honoring Lila Gleitman," Barbara Landau (Topic Editor). For a full listing of topic papers, see http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview

Morphology and Memory: Toward an Integrated Theory

Ray Jackendoff,[a,b] Jenny Audring[c]

[a]*Center for Cognitive Studies, Tufts University*
[b]*Department of Brain and Cognitive Sciences, MIT*
[c]*Centre for Linguistics, Leiden University*

Received 2 March 2017; received in revised form 6 February 2018; accepted 16 February 2018

Abstract

Framed in psychological terms, the basic question of linguistic theory is what is stored in memory, and in what form. Traditionally, what is stored is divided into grammar and lexicon, where grammar contains the rules and the lexicon is an unstructured list of exceptions. We develop an alternative view in which rules of grammar are simply lexical items that contain variables, and in which rules have two functions. In their *generative* function, they are used to build novel structures, just as in traditional generative linguistics. In their *relational* function, they capture generalizations over stored items in the lexicon, a role not seriously explored in traditional linguistic theory. The result is a highly structured lexicon with rich patterns among stored items. We further explore the possibility that this sort of structuring is not specific to language, but appears in other cognitive domains as well, such as the structure of physical objects, of music, and of geographical and social knowledge. The differences among cognitive domains do not lie in this overall texture, but in the materials over which stored relations are defined. The challenge is to develop theories of representation in these other domains comparable to that for language.

Keywords: Morphology; Lexicon; Memory; Words and rules

We are grateful for the opportunity to contribute to this special issue in honor of Lila Gleitman and her Rumelhart Prize. We are in awe of Lila for her splendid taste in restaurants, cheeses, and jokes, but more (of course) for her spectacular mentoring of generations of groundbreaking researchers, and even more for many decades of her own

Correspondence should be sent to Ray Jackendoff, Center for Cognitive Studies, Tufts University, Medford, MA 02155. E-mail: ray.jackendoff@tufts.edu

stunning work. And behind it all is her abiding passion for the deep questions posed by the study of language. The present paper is an attempt to channel that questioning spirit.

1. "Knowledge of language"

The issue we are concerned with here is morphology, the grammar of words. We approach it with a number of goals in mind. The narrowest is to sketch a theory of morphology, Relational Morphology, that is consistent with the Parallel Architecture framework for linguistic theory (Culicover & Jackendoff, 2005; Jackendoff, 1997, 2002). More broadly, we wish to situate morphology in an integrated theory of language in which phonology, morphology, syntax, and semantics fit together naturally. This task is not just a matter of grafting morphology onto existing theory: It turns out that considerations from morphology affect the way we think about syntax as well. And ultimately, we aspire to a still more ambitious goal: figuring out how the components of the language faculty fit naturally into the rest of the mind.[1]

We do not pretend to be completely original here. Many aspects of our approach have been in the literature and "in the air" for a long time. What we think is relatively original is our synthesis of these ideas into a coherent whole.

Like Mainstream Generative Grammar (MGG: Chomsky, 1965, 1981, 1995) and many other approaches to linguistic structure (especially Construction Grammar: Goldberg, 1995, 2006)—and in concurrence with all research in language processing and language acquisition—we take language to be a mental phenomenon. Indeed, we are somewhat more concrete about our mentalism than is MGG, in that we construe Chomsky's (1986) term "knowledge of language" very literally. From the point of view of psychology, "knowledge" must somehow be in the brain, stored in memory. Hence, the study of knowledge of language amounts to investigating what information is stored in long-term memory, how it is used, and how it comes to be there in the first place. With this framing, the fundamental questions of linguistic theory can be restated like this:

1. What linguistic elements does a speaker store in memory, and, crucially, in what form?
2. How are these elements combined online to create novel utterances?
3. How are these elements acquired?

The second of these, the creation of novel utterances, has been the central concern of generative grammar since its inception (Chomsky, 1957, 1959). The primary goal has been to characterize Humboldt's (1836/1999) "infinite use of finite means"—the generation of an unlimited variety of forms from basic pieces stored in the lexicon. This emphasis continues up to the present; for example Berwick and Chomsky (2016: 2) speak of the "Basic Property of language: that a language is a finite computational system yielding an infinity of expressions." We wholeheartedly endorse the importance of this property of language. But, as will be seen, we take issue with the degree of primacy it has been

afforded, to the detriment of concern with the "finite means" from which the infinity of expressions is constructed.

Among the finite means, of course, is the lexicon. A standard assumption, inherited from traditional grammar and the structuralists (Aronoff, 1994, traces it all the way back to the 17th century), is that the lexicon is the place where words live. It is strictly distinguished from the grammar, which consists of a collection of productive, general rules. An ancillary assumption is that the lexicon is the locus of idiosyncrasy, and the grammar captures all linguistically significant regularity. Since it is the rules that are responsible for building the unlimited variety of utterances, the lexicon has tended to take a back seat.

We take issue with these assumptions, as do other frameworks, particularly Cognitive Grammar (Langacker, 1987); Construction Grammar (Croft, 2001; Fillmore, Kay, & O'Connor, 1988; Goldberg, 1995, 2006; Hoffman & Trousdale, 2013), and Construction Morphology (Booij, 2010). In these approaches as well as ours, rules turn out actually to be treated as lexical items, as will be explained in Section 2.

This is where morphology comes in. Morphological phenomena hit us in the face with their intricate mixture of regularity and idiosyncrasy. Unlike syntactic structures, a significant proportion of morphologically complex words (for instance, *substantial*) cannot be built directly from smaller pieces (such as *substance* and *-al*). Rather, these words have to be learned and stored. At the same time, words like *construction*, *wishful*, and *glorious* have internal structure, and they are not entirely "exceptional." Psycholinguistic research has in fact shown that morphological structure plays a role in language processing—it is not just a matter of linguists reading too much into the data (Amenta & Crepaldi, 2012; Baayen, Dijkstra, & Schreuder, 1997; Bybee, 1995; Clahsen, 1999; Hay & Baayen, 2005; Levelt, 1989; Moscoso del Prado Martin, Bertain, Häikiö, Schreuder, & Baayen, 2004; Nooteboom, Weerman, & Wijnen, 2002; Pinker, 1999; and hosts of others).

Returning to the fundamental questions above, we will conclude that linguistic theory has to be concerned not just with the infinite use—the active generation of productive regular patterns—but also with characterizing the finite means: the structure of lexical items stored in memory, the properties they share with other lexical items, and the nature of their idiosyncrasies. Nor will this conclusion be confined to morphology. We will show that the same issues also arise in syntax, although less pervasively, hence easier to overlook (Culicover et al., 2017).

A reviewer has objected to this focus on idiosyncrasy, on the grounds that mainstream generative grammar has always understood exceptions to be theoretically irrelevant, as explained for instance in *Aspects* (Chomsky, 1965: 218):

> It must be borne in mind that the general rules of a grammar are not invalidated by the existence of exceptions. . . . It is for this reason that the discovery of peculiarities and exceptions (which are rarely lacking, in a system of the complexity of a natural language) is generally so unrewarding and, in itself, has so little importance for the study of the grammatical structure of the language in question . . .

This gives the impression that the "exceptions" are a minor matter. But in morphology they are ubiquitous: They are less like Sapir's (1921: 39) "leak in the grammar" and more like a torrent. A comprehensive theory of knowledge of language cannot simply declare them "unrewarding" and "unimportant." It must include an account of the irregularities and the subregularities among them.

The quote above continues with a codicil:

> . . .unless, of course, it [i.e. the discovery of peculiarities and exceptions – RJ/JA] leads to the discovery of deeper generalizations.

If, as advised, one disregards the study of irregularities, one will never find out if they lead to any deeper generalizations. They might or they might not. We propose to show here that there are indeed deeper generalizations which encompass both the productive and the irregular facets of language. But in order to achieve these generalizations, one has to rethink some of the Received Wisdom about the relation of lexicon and grammar.

2. Lexical items, links, and schemas

To situate our sketch of morphology, we quickly review the Parallel Architecture, developed in detail in Jackendoff (1997, 2002). Its basic tenet is that linguistic structure is determined by three independent generative systems—phonology, syntax, and semantics—plus the linkages or interfaces between them. (1) sketches the overall architecture; variants in a similar spirit appear in Autolexical Syntax (Sadock, 1991), Lexical-Functional Syntax (Bresnan, 1982, 2001), and Role and Reference Grammar (Van Valin & LaPolla, 1997).[2]

(1) Phonological structures ⟷ Syntactic structures ⟷ Conceptual (semantic) structures

A well-formed sentence has well-formed structures in each of the three domains, plus well-formed links among the structures.

Similar linkages are to be found between other domains of mental representation. For instance, in order to be able to speak about what one sees, there has to be an interface that links linguistic representations to representations in the visual system. Likewise, in order to coordinate visual shape recognition with haptic shape recognition, there must be a linkage between representations created in the visual system and those created in the haptic system, by virtue of which seen shapes and felt shapes can be judged as equivalent. Hence, the PA conception of the language faculty fits naturally into a larger perspective on how distinct cognitive domains interact with each other (for more discussion, see Jackendoff, 2011, 2015).

This conception of the language faculty contrasts with MGG, in which syntax is the only "generative engine," and in which phonological and semantic structures are derived

from syntactic structure. Such a conception does not tell us how to make contact with the rest of the mind; this has never been a concern of MGG. In particular, visual representations cannot be derived from syntax or vice versa; nor can they be derived from haptic representations or vice versa. Even internal to language, the assumption of syntactic primacy is problematic. Culicover and Jackendoff (2005, especially chapters 2 and 3) show how much of the complexity of MGG syntactic structure stems from the assumption that semantic complexity is to be accounted for in terms of syntax. They further show that a Parallel Architecture approach leads to syntactic structures that are not only simpler but also more empirically adequate. We set these arguments aside here and simply assume the PA framework.

The reader will have noticed that (1) does not mention a lexical component in the architecture. To see why not, consider what a word looks like in this framework. (2) notates the structure of the word *pig*.

(2) Semantics: PIG_1
Morphosyntax: N_1
Phonology: /pɪg/$_1$

The word consists of a piece of structure in each of the three components. Its semantics, notated as PIG, is whatever the word means; its morphosyntactic features identify it as a noun; and its phonological structure specifies how it is pronounced. In addition, the three structures share coindex 1. This coindex marks ***interface links*** between the three components—an explicit association between them in long-term memory. Thereby, the word constitutes a small part of the interface.[3] (2) says that if you encounter the phonology /pɪg/, you are entitled to posit the associated meaning PIG with the corresponding syntactic feature Noun. Similarly, if you wish to express the meaning PIG, you can do so with a Noun pronounced /pɪg/.

So far this is fairly boring. It gets more interesting when we look at a morphologically complex word like *piggish*. (3) shows its structure.

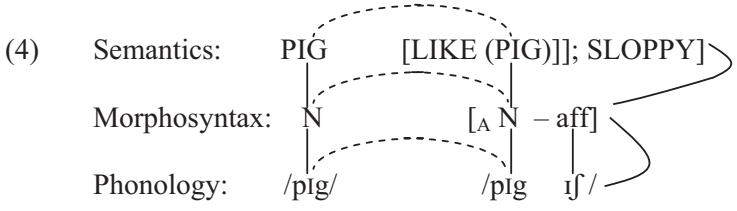
(3) Semantics: $[LIKE_3\ (PIG_1);\ SLOPPY,\ GREEDY]_2$
Morphosyntax: $[_A\ N_1\ \ aff_3\ ]_2$
Phonology: /pɪg$_1$ ɪʃ$_3$ /$_2$

Coindex 2 notates interface links among the three components of *piggish*, in the same way that coindex 1 in (2) connects the components of *pig*. Coindex 3 notates the contribution of the suffix to all three levels: It means "like an *x*"; morphosyntactically it is an affix; and it is pronounced /ɪʃ/. Coindex 1 is the interesting one. Within the entry of *piggish*, it functions as an interface link that connects the three components of the base. But in addition, by virtue of being the same as the coindex of *pig* in (2), it also functions as a ***relational link*** between corresponding components of the two words.

Notice, however, that *piggish* cannot be derived from *pig*, as its meaning is more specific than merely "like a pig." Its full meaning is something like "sloppy and/or greedy like a pig"; it does not mean, for example, "pink like a pig." Hence, it must be learned and stored in long-term memory. This means that the relational links notated by coindex 1 are relations within the lexicon, and therefore that *piggish* is not "atomic."[4]

To be a bit more explicit about the status of the coindices: They are to be understood as marking the ends of association lines between pieces of structure in the lexicon. We could instead use association lines to notate *pig*, *piggish*, and their connection, as in (4). The solid lines are interface links, which connect different levels within the same word; the dashed lines are relational links, which connect the same levels of different words.[5]

(4)



However visually appealing this notation might be, it rapidly degenerates into spaghetti when one tries to do anything at all complicated. We therefore retain the coindexing notation, despite its somewhat less iconic character.

In the case of *piggish*, the relation between the three levels of structure is one-to-one: for each part in phonology, there are corresponding parts in morphosyntax and semantics, and vice versa. But this is not always the case. Consider the plural of *sheep*, namely *sheep*, notated in (5).

(5)

Semantics: $[\text{PLUR}_6\ (\text{SHEEP}_4)\ ]_5$

Morphosyntax: $[_N\ N_4\ \ PL_6\ ]_5$

Phonology: $/ʃip/_{4,5}$

Here, morphosyntactic plural is necessary to govern determiner and verb agreement (*those sheep are*... vs. *that sheep is*...). It is linked to the plural function in semantics by coindex 6. However, coindex 6 does not link to anything in phonology. Rather, the plural is pronounced the same as its base, as marked by the dual coindexation in the phonology. Thus the present approach does not need a "zero morpheme" in phonology; zero morphology is simply a piece of morphosyntax unlinked to anything in phonology. (For the treatment of many other noncanonical morphological relations in terms of this formalism, see our forthcoming book, *The Texture of the Lexicon*.)

Next we wish to express the parallelism among all the *-ish* words—*piggish*, *childish*, *foolish*, *thuggish*, and so on. In order to do this, we introduce a ***schema*** for the *-ish* suffix. (6) gives the structure of two more *-ish* words, and (7) is the schema.[6] Informally, (7)

says that an adjective ending in *-ish* can mean "like X," where the meaning "X" is expressed by a noun whose phonology precedes *-ish*.

(6) a. Semantics: $[\text{LIKE}_3\,(\text{CHILD}_7);\ \text{SILLY, IMMATURE}]_8$
Morphosyntax: $[_A\ N_7\ \ \text{aff}_3\ ]_8$
Phonology: /tʃɑɪld$_7$ ɪʃ$_3$/$_8$

b. Semantics: $[\text{LIKE}_3\,(\text{FOOL}_9)]_{10}$
Morphosyntax: $[_A\ N_9\ \ \text{aff}_3\ ]_{10}$
Phonology: /fu:l$_9$ ɪʃ$_3$/$_{10}$

(7) Semantics: $[\text{LIKE}_3\,(X_x)]_y$
Morphosyntax: $[_A\ N_x\ \text{aff}_3\ ]_y$
Phonology: / …$_x$ ɪʃ$_3$ /$_y$

The interpretation of the notation in (7) is as follows.

1. The components of the affix in (7) are linked by coindex 3. Coindex 3 appears in all the instances of the schema in (3) and (6), hence it serves as a relational index identifying *-ish* as the affix that they all share.
2. (7) contains variables in semantics, syntax, and phonology. These are linked together by the ***variable coindex*** *x*. This coindex can form a relational link with any word that has the same pattern of structure. For example, *x* in (7) forms a relational link with coindex 1 in *piggish*, with coindex 7 in *childish*, and with coindex 9 in *foolish*.
3. The three levels of the schema as a whole are linked by the variable coindex *y*. This too can form a relational link to anything with the same pattern of structure. For instance, *y* in (7) forms a relational link with coindex 2 in *piggish*, coindex 8 in *childish*, and coindex 10 in *foolish*.
4. SLOPPY, GREEDY, SILLY, and IMMATURE are not connected with anything in the schema; these are the idiosyncratic parts of the word meanings.

Hence we might say that coindex 3 notates the ***same*** structure in the schema and its instances, while *x* and *y* notate ***equivalent*** or ***parallel*** structure.

This being the lexicon, there are all sorts of partial resemblances among words ending in *-ish*. *Bookish* and *feverish* share coindex 3 on phonology and morphosyntax, but they do not mean "like a book" or "like a fever." Adjectives such as *oldish* and *warmish* have as their base an adjective rather than a noun, and they mean something like X TO SOME DEGREE. *Squeamish* and *brackish* are adjectives, but their bases are not words at all (an important situation to which we return); still, they can be identified as containing the suffix *-ish* in morphosyntax and phonology. Finally, words such as *radish*, *rubbish, furnish*, and *vanish* also end in *-ish*, but they are not even adjectives, so they share nothing with

words with the *-ish* suffix, aside from accidental phonology. Each of these cases is "exceptional" in its own way, and the deviations can be expressed in terms of the presence or absence of particular relational links to schema (7).

3. Schemas versus rules

(7) encodes a "word formation rule" in terms of a declarative schema—a description of a possible structure in English. As mentioned above, it can be informally paraphrased as (8a). This general characterization is shared by other constraint-based frameworks such as LFG, HPSG, Cognitive Grammar, Construction Grammar, and Construction Morphology. In contrast, the rules of MGG are procedures like (8b); they have an "input" and "output," and they apply in sequence to perform derivations.[7]

(8) a. *Declarative formulation:*

> An adjective ending in *-ish* can mean 'like X', where the meaning 'X' is expressed by a noun whose phonology precedes *-ish*.

b. *Procedural formulation:*

> Take a noun that means 'X' as input, and add *-ish* to its phonology, outputting an adjective that means 'like X.'

At first glance, these two formulations might look like notational variants, and one might wonder why it matters which of them one chooses. Why should one should give up rules in favor of schemas? How are they different?

The most important difference is that a schema like (7) is in the very same format as words: it consists of pieces of semantics, morphosyntax, and phonology, connected to each other by interface links and connected to their instances by relational links. This means that, as promised, it is no longer necessary to separate the theoretical constructs of "lexicon" and "grammar"; we can reduce the major components of language from two to one. The distinction between words and rules is gradient and emerges from the structure of lexical items themselves. A structure in the lexicon is more "word-like" if its content is completely filled out (e.g., *pig*), and more "rule-like" if it contains variables (e.g., schema (7)). The most rule-like items consist entirely of variables, as we will see below. Such a gradient conception of words and rules is impossible under the traditional view in which words and rules are of entirely different format and occupy different "places" in the mind.

The schema format is also natural for encoding the subcategorization of a transitive verb like *devour*. Instead of assigning to its lexical entry a subcategorization feature like the traditional (9a) or a feature that marks case on the object, the entry can be stated in schema form, for a first approximation like (9b).

(9) a. [+ __ NP]

b. Semantics: $[\text{DEVOUR}_{11}\ (\text{Agent: X, Patient: } Y_y)]_z$
Syntax: $[_{VP}\ V_{11}\ NP_y\ ]_z$
Phonology: /dəvawr$_{11}$ …$_y$ /$_z$

(9b) says that a VP whose verb is pronounced /devawr/ contains a variable (i.e., an open NP slot), which is linked to the Patient argument of the semantic function DEVOUR (X,Y). If this NP is absent, the VP is ill-formed. This is exactly what a subcategorization feature is meant to say. This way of formulating it, however, recognizes that the verb *devour* is somewhat rule-like, in that it contains a variable in its lexical entry, unlike *pig*, which lacks variables altogether.

This folding of grammatical rules into the lexicon does not concern just words. It also extends readily to syntactic phrases. For instance, the phrase structure rule for transitive VPs, instead of being stated as a procedure to "expand" or "rewrite" the node VP, as in (10a), can be recast as a schema that has only a syntactic component, as in (10b). Since (10b) lacks semantics and phonology, it also lacks interface links. Such schemas are what Janet Fodor, 1998 calls "treelets"; they are also the fundamental building blocks of Tree-Adjoining Grammar (Joshi, 1987).[8]

(10) a. *Procedural formulation*:
VP → V – NP

b. *Declarative formulation:*
Syntax: $[_{VP}\ V\ \ NP]$

One of main themes of LFG, HPSG, and Simpler Syntax (Culicover & Jackendoff, 2005) is to show that a declarative account of syntactic patterns can be carried through all the standard syntactic phenomena such as passive, raising, control, and long-distance dependencies. (See also Pullum, 2007, who contrasts the characterization of grammatical patterns as descriptions of structure versus procedures for building structure.)

This interpretation of phrase structure rules presents a possibility not available to the standard conception: some syntactic structures might be more "word-like," by virtue of being intrinsically associated with characteristic meanings. This possibility has been a preoccupation of Construction Grammar, back to its founding documents (Fillmore et al., 1988). An example is the demonstration by Gleitman, Gleitman, Miller, and Ostrin (1996) that symmetrical predicates such as *similar* behave asymmetrically in the configuration *X is similar to Y*, not because they are inherently asymmetrical, but because subject position imposes a figural perspective on X, against which Y is the standard of comparison. That is, the asymmetry arises from semantic structure that is linked directly with

syntactic structure, not with particular words (see also Landau & Gleitman, 2015). This sort of generalization is difficult to capture in a traditional architecture, in which syntactic structures themselves can carry no meaning.

The fact that schemas are in the same format as words has another advantage: it makes an account of their acquisition formally straightforward. Every theory of acquisition needs a mechanism that detects parallelisms among stored items and uses them as the basis for positing a tentative rule. In a traditional architecture, such a hypothesized rule is a procedure—an entirely different sort of entity from a lexical item—so the acquisition process must so to speak cross a transcendental gap from word format to rule format.[9]

A mechanism for forming schemas, in contrast, simply creates a new lexical item—the schema—by copying the parts of the parallel instances that are the same, and replacing with variables the parts that are different among the instances. For instance, the *-ish* schema (7) has constants coindexed 3 where *piggish*, *childish*, and *foolish* are the same, and variables where they are different. In other words, the schema can be derived directly from the lexical items that motivate it. The same sort of procedure has been advocated for phrasal schemas such as (10) by, among others, Tomasello (2003) and Culicover and Nowak (2003). We take this to be an immediate formal advantage of schemas, and another respect in which they are not simply a notational variant of traditional rules. The next section offers further differences between schemas and rules.

To sum up this section: The distinction between lexicon and grammar is so deeply ingrained that it may be difficult to conceive of abandoning it. However, we have already begun to see some advantages in doing so. First, instead of having two components of grammar with entirely different formats, we have only one. The traditional distinction between words and rules surfaces instead as the absence or presence of variables in lexical items, a more flexible distinction. Second, this formalization allows for meaningful constructions: pieces of syntax that, just like words, are associated directly with meaning, a possibility denied in MGG but amply supported by evidence from Construction Grammar. Third, using schemas as the central construct for capturing both morphological and syntactic patterns also allows us to establish a continuity between morphosyntax and phrasal syntax, while still preserving the differences between them. Fourth, the acquisition of schemas is formally simplified. Thus, already in several respects we are on the way to a better integrated linguistic theory.

4. Problems with rules

To further demonstrate that schemas are not simply a reformulation of traditional rules, we now turn to some problems faced by traditional rules, and their solution in terms of schemas—more respects in which schemas are more flexible than rules.

Rules such as the VP rule (10a) (or any other procedural rule such as Merge) are taken to be the source of creative combinatoriality. They combine items stored in the lexicon (say, *forget*, *that*, and *fox*) and produce structures such as [$_{VP}$ *forget* [$_{NP}$ *that fox*]] which

are *not* stored in the lexicon. However (following Weinreich, 1969/1980), consider idioms like *shoot the breeze*, *spill the beans*, *chew the fat*, and the ever-popular *kick the bucket*. They have the syntactic structure of ordinary VPs, and their verbs inflect according to normal patterns. In particular, the past tense of *shoot the breeze* is *shot the breeze*, just like the ordinary verb *shoot*. So these phrases should be generated by the VP rule. Yet they cannot be generated by rule, because their meanings are wrong—they have nothing to do with shooting, breezes, and so on. And because of their noncanonical meanings, they have to be stored in the lexicon, although they are not "syntactic atoms." Thus there is a paradox: Given their canonical syntactic structure, these items should be generated by the VP rule. But there is no way for the VP rule to generate them, because they are listed in the lexicon, and the output of the VP rule is by definition outside the lexicon. It should also be noted that idioms are far from isolated "exceptions" that can be safely disregarded: there are probably as many idioms in English as adjectives.[10]

A language user's knowledge also includes a plethora (probably tens of thousands) of conventionalized phrases and sentence fragments such as *you know what I mean* and *I'm terribly sorry* that mean exactly what they should, but are known to be "the right way" to say things in the language (Christiansen & Arnon, 2017; Culicover et al., 2017; Jackendoff, 1997; Pawley & Syder, 1983). Again, if they are stored in the lexicon, then the rules of grammar cannot generate them.

Morphology presents many such cases. Here are four.

1. A word such as *football* has canonical compound structure, and by virtue of this one can surmise that it denotes a kind of ball, and its meaning also has something to do with feet. But its meaning involves huge idiosyncrasy—not the least of which is that it can also denote a game that *uses* a football. Thus, its meaning is a mixture of rule-governed and idiosyncratic properties.
2. Words with irregular phonological alternations pose similar problems, but in the phonological component. For instance, an English speaker has to learn and store the fact that the past of *sing* is *sang* but that of *sting* is not **stang* but *stung*, and that of *bring* is *brought*. There is an overall tendency for monosyllabic verbs ending in /ŋ/ to be irregular (Bybee & Slobin, 1982), but each of these examples falls under a different pattern, and as a result, the lexicon has to list the irregular past tenses alongside the corresponding present tenses. But if these forms are stored, there is no way for a phonological rule, by definition outside the lexicon, to change present tense forms of irregular verbs into past tense forms.
3. There are families of words with perfectly predictable phonology and meaning, but which have to be listed anyway. An example is the family of deadjectival verbs with an *-en* suffix, such as *widen*, *stiffen*, and *sicken*. These are all built on monosyllabic adjectives that end in an obstruent, and they all mean "(cause to) become (more) X." However, some adjectives satisfy these criteria perfectly but do not form *-en* verbs, for instance **louden*, **safen*, and **crispen*.[11] The simplest way to distinguish the forms that exist from those that do not is to list the forms that exist. But that means that the regularity of the pattern cannot be captured by a rule along

the lines of "Add *-en* to an adjective that means 'X' to create a new word that means 'become X'." The output of such a rule is by definition outside the lexicon, but the derived words have to be inside the lexicon.

4. There are lots of morphologically complex words that have an identifiable affix, but their base is not a lexical item, for example, *commotion* (**commote*) and *impetuous* (**impet*)—compare with *devotion* (*devote*) and *incestuous* (*incest*). *Impetuous*, for instance, cannot be constructed by a rule that says "to form an adjective, add *-ous* to a noun," since there is no noun to add *-ous* to.

In addition, experimental evidence (e.g., Alegre & Gordon, 1999; Baayen et al., 1997) shows that even many regular composite forms are stored. But again, under MGG assumptions, if they are stored, they cannot be derived by rule. Hence, their regularity is unaccounted for.

To sum up these examples, we see that rules, under the standard conception, both overgenerate and undergenerate. They overgenerate in cases like the A-*en* verbs, where any rule that derives *widen* from *wide* cannot be prevented (without further machinery) from deriving **crispen* from *crisp*. And they undergenerate in cases like *football*, where they cannot derive all the details of meaning, and in cases like *squeamish*, where there is nothing to derive it from.

A widely adopted solution to these difficulties was Chomsky's (1970) Lexicalist Hypothesis. The idea was that regular, predictable patterns belong to syntax, whereas idiosyncratic patterns such as these are encoded by "lexical redundancy rules" (or simply "lexical rules"), situated in the lexicon. Hence, if an item shows any amount of idiosyncrasy, it belongs in the lexicon, and any regularity it may display is a consequence of lexical rules. The Lexicalist Hypothesis was widely adopted, for example, in Wasow's (1977) influential paper distinguishing "grammatical passive" from "lexical passive." The existence of a separate component of lexical rules became a basic assumption in HPSG and LFG ("Lexical Integrity") as well as in Lexical Phonology (Kiparsky, 1982; Mohanan, 1986).[12]

The effect of the Lexicalist Hypothesis is to further bolster the strict distinction between lexicon and grammar. It removes from the syntactic component all traces of impurity—a sort of grammatical apartheid. Yet consider what is being excluded. The "exceptional" cases are not negligible, mere "leaks" in the grammar. Thousands of compounds like *football* are lexicalized; about 200 of the approximately 600 adjectives ending in *-ous* are like *impetuous* in having a nonlexical base (e.g., *invidious*, *insidious*, *fastidious*), and the same goes for about 700 of the approximately 900 verbs ending in *-ate* (e.g., *instigate*, *irrigate*, *mitigate*). Moreover, none of these examples are entirely exceptional: there are clearly patterns or subregularities in play.

Furthermore, the Lexicalist Hypothesis still does not answer how the grammar generates *shoot the breeze*. Apparently, the solution is to add a VP rule to the component of lexical rules. But this too is odd: Why should there be two identical rules, one in the grammar and one in the lexicon?

5. The solution: Two functions of schemas

Facing this last problem, Jackendoff (1975) proposes that the phrase structure rules can also act as lexical redundancy rules. Acting as a phrase structure rule, the VP rule generates novel phrases like *examine the pizza*; as a lexical redundancy rule, it accounts for stored VPs such as *shoot the breeze*. Aronoff (1976) makes a similar proposal for morphology: his Word Formation Rules act both as lexical rules and as generative rules that can form new words like, say, *Trumpish*. However, within the theoretical outlook of the time, these proposals were not entirely cogent: How could a rule be in the grammar and still act like it is in the lexicon—or vice versa? As a result, these proposals never really caught on.

However, they make more sense in the context of Relational Morphology and other constructionist approaches. For one thing, syntactic rules are *in* the lexicon, so they are in the same "place" as lexical redundancy rules, not in an entirely different component of grammar. Moreover, both sorts of rules are stated in schema format, which in turn is the same format as words. Hence, we are not just sneaking into the lexicon entirely new entities, namely two sorts of rules. We are just adding variables and variable indices (which, as we have seen, we need anyway to state subcategorization features on words).

The heart of our proposal is this: unlike traditional rules, a schema can function in two distinct roles, the ***generative*** role and the ***relational*** role.

In their ***generative*** role, schemas are used to build up novel composite expressions by unifying the schema's variables with other material—either words or other structures. This role corresponds to that of traditional rules. Its procedural nature comes from the general procedure of Unification (Shieber, 1986) that actively assembles pieces into larger structures, in particular instantiating variables of schemas.[13]

In their ***relational*** role, schemas capture generalizations among items stored in the lexicon. The literature refers to this function as "motivation" (e.g., Booij, 2017; Goldberg, 1995, 2006) or "inheritance" (e.g., Pollard & Sag, 1994; Goldberg again); it is essentially the function attributed to lexical redundancy rules. In this role, a schema does not have to exhaustively specify the contents of the items being related. There is latitude for idiosyncrasy such as GREEDY in *piggish*; the schema only specifies predictable aspects of its instances. (See our forthcoming book, *The Texture of the Lexicon*, for discussion of the mechanisms of motivation/inheritance in terms of relational links.)

The insight of Jackendoff (1975) and Aronoff (1976) can now be implemented directly:

The same schema can function in both the generative role and the relational role.

For example, compounding is highly productive—one encounters and coins novel compounds all the time.[14] Hence the compound schema has to function in a generative role to create these new structures. But at the same time, as mentioned above, there are thousands of lexicalized compounds like *football* that conform to the compound schema but have idiosyncratic interpretations. These too fall under the compound schema, but in its relational role.

Similarly, the schema for the English plural can apply to generate novel forms with novel or rare bases such as *wugs*, *fendles*, and *coelacanths*. At the same time, if we are to trust the psycholinguistics, the lexicon stores lots of frequent and completely regular forms such as *cats*, and does not need to generate them online. Nevertheless, the schema for the regular plural still applies to these stored items. It is the same schema, the same regularity, but realized via the relational rather than the generative role of the plural schema.

We note also that there are idiomatic items that contain regular plurals, such as *raining cats and dogs*, *make amends*, *shake hands*, *goings-on*, and *odds and ends*. Since such idioms have to be stored, the regular plurals within them have to be stored as well. Yet the plural schema, again in its relational role, takes account of their regularity.

Admitting these dual roles of schemas solves the problem of undergeneration. When an idiosyncratic item has to be stored, but still appears to fall partly under a regular pattern, the schema in question is being used relationally rather than generatively, and hence is not required to predict everything about every instance.

Turning to the problem of overgeneration: This can be solved by proposing that some schemas function *only* in the relational role. Consider again the deadjectival verbs *harden*, *darken*, *widen*, *stiffen*, and so on. The operative schema is (11).

(11) Semantics: $[\text{BECOME}_{12}\ (\text{PROPERTY}_x)]_y$
Morphosyntax: $[_V\ A_x\ \text{aff}_{12}\ ]_y$
Phonology: $/\ \ldots_x\ \text{ən}_{12}\ /_y$

Despite the apparent regularity of this pattern, one has to learn the A-*en* verbs individually (as well as idiosyncrasies of individual instances, such as *lengthen* and *embolden*). There is no temptation to accept (or coin) forms such as **safen*, **louden*, and **crispen*. This is exactly what we would predict if (11) can be used only in its relational role: It is nonproductive and cannot build new items.

A similar approach applies, of course, to all the little patterns of irregular English past tenses such as *sing/sang*, *drink/drank*, and so on. Each pattern has a schema, but it only functions relationally. Hence all the instances of the pattern have to be stored in the lexicon.

In short, the patterns for which a productive rule would overgenerate instead fall under *non*productive schemas. And in fact, the bulk of morphological patterns in English are like this; productive schemas such as the regular plural, regular past, and $[_N$ A-*ness*] (as in *happiness* and *picturesqueness*) are in a distinct minority.

Extending this approach to syntax, we find that it solves the problem of the syntax of idioms. As always, the transitive VP schema (10b) applies generatively to create novel VPs such as *discard the printer*. But it also applies relationally to capture the generalization inherent in listed instances such as *shoot the breeze*. As with the plural, it is the very same schema, just used differently. The same goes for the syntax of conventionalized phrases such as *I'm terribly sorry*, paralleling the storage of morphologically regular words.

Exploring parallels between syntax and morphology a little further, let us ask: Could there be *syntactic* schemas that, like many morphological schemas, function only relationally and are therefore restricted to instances listed in the lexicon? This is of course impossible in the MGG canon. But indeed, if we poke around in the odd corners of syntax, we do find such exotic beasts. One example is the NPN construction, as in *day after day* (Jackendoff, 2008). The choice of doubled noun is free in this construction, but the choice of preposition is confined to five or six choices, listed in (12a), each with its own idiosyncratic semantics. The prepositions in (12b), although potentially meaningful, are impossible.[15]

(12) a. day after day
week by week
face to face
dollar for dollar
book (up)on book

b. *person from person
*gun beside gun

Another case is the peculiar little determiner pattern in (13).

(13) a. what/such/quite/hardly/many/nary a book
b. how/so/too/as/this/that lovely a story
c.*where a city
d.*very lovely a story

In both of these constructions, there is a generalization, but the instances have to be learned one by one. They cannot be coined on the spot.[16] Again, these are the symptoms of a schema that can only be used relationally—but this time in syntax. These constructions may be off the beaten track and "exceptional," but they are not bereft of pattern, and these patterns are part of our knowledge of English. So here, an innovation in morphology offers an approach to a relatively neglected problem in syntax.

The question arises of how a productive schema, which functions both generatively and relationally, differs from a nonproductive schema, which functions only relationally. A simple solution is to mark the schema for productivity (whether this is a categorical difference or a gradient one—we are not committed). This is the solution advocated by Booij (2010); and also in a sense by Lakoff (1970), whose "major rules" apply across the board and whose "minor rules" apply only to listed instances. A more delicate approach marks productivity not on the schema as a whole but on its variables. This allows for the possibility that a schema might have one nonproductive variable and one productive

variable. Such a possibility is realized in (12), where the choice of preposition is nonproductive and the instances have to be listed, but the choice of doubled noun is completely free (with the exception of the *face to face* case, which admits only listed nouns). Similarly in (13b), the choice of degree word is restricted to listed options but the choice of adjective is entirely free.

The upshot, then, is that instead of a grammar in which there is a strong distinction between "grammatical rules" and "lexical rules," we have a grammar where the difference between the two phenomena is reduced to a choice of diacritic on a variable—a substantial reduction in the overall complexity of grammatical theory. We take this to be a "deeper generalization" of the sort grudgingly suggested in Chomsky's codicil in section 1.

6. Productivity and acquisition

To cement the continuity between productive ("grammatical") and nonproductive ("lexical") patterns, we turn again to acquisition. An essential part of this process, of course, is constructing (or discovering) the productive rules/schemas of the language, on the basis of primary linguistic input. As suggested in section 3, the learner's procedure must involve observing some collection of words with similar phonological and semantic structure, and formulating a hypothesis about the general pattern they instantiate (Culicover & Nowak, 2003; Tomasello, 2003).

Crucially, learners have no way of knowing in advance whether a pattern they observe will prove to be fully productive or not. In a theory that distinguishes grammatical rules from lexical rules, which kind of rule should the learner hypothesize? Suppose the learner posits a grammatical rule. Then, if the rule fails the criteria for productivity (whatever these criteria may be—see, for instance, Baayen & Lieber, 1991; Haspelmath & Sims, 2013; Yang, 2016; O'Donnell, 2015), it must be either expunged or "converted" into a lexical rule, that is, moved into a different component of the grammar. Alternatively, if the learner posits a lexical rule, and sufficient evidence arrives to deem it productive, then it has to be "converted" in the opposite direction.

In contrast, in a schema-theoretic approach, productive and nonproductive schemas are in exactly the same format. If a learner extracts a pattern as a tentative schema, it might or might not be a productive one, and the next job is to determine whether this schema is productive or not. This is not a transcendental distinction between a rule "in the grammar" and one "in the lexicon," as in the Lexicalist Hypothesis, or between a rule versus no rule, as in Pinker's (1999) "words and rules" approach. It is just a matter of determining the proper diacritic on the schema's variable. Formally, this is a relatively small and local issue.

Moreover, if a schema turns out to be nonproductive, this does not mean it is flat-out wrong. The observed pattern among the observed instances may still remain valid. And if it so happens that a schema is found to be productive, it still does not relinquish its status as a lexical redundancy rule. Rather, as suggested in the previous section, it still retains the function of capturing generalizations among stored items.

Nonproductive schemas can of course be useful as an aid in acquisition of new instances. When one encounters a new word, one presumably seeks patterns into which it fits. Without schemas, there are endless ways a new word can be similar to existing words, along one dimension with one word (e.g., initial syllable), another dimension with another word (meaning), a third dimension with a third word (final syllable), and so on. A schema codifies dimensions of similarity that have been found significant, in effect "precompiling" the similarities among all its instances.

7. Reframing the goals of linguistic theory

At this point we have arrived at (or perhaps stumbled into) a radical conclusion. We have encountered two kinds of schemas: those that can be used both generatively and relationally, and those that can only be used relationally. It is logically possible that there are also schemas that can be used only generatively. Indeed, that is the way traditional rules are conceived. But in fact there cannot be such schemas. Why not? The reason is that anything that can be generated online can then be committed to memory. And once it is committed to memory, the schemas originally involved in generating it now have to account for it relationally.

Therefore, we conclude that *productive* schemas both generate new expressions and motivate stored items, while *nonproductive* schemas just motivate stored items. The two kinds of schemas are in the same format. The difference is only in which roles they can play.

This state of affairs can be reframed in an interesting way: *All* schemas can be used relationally. Productive schemas are just like nonproductive schemas, except that they *also* allow free online composition. That is, we can think of productive schemas as a special subset of the much larger class of linguistic patterns—they are ordinary schemas that have "gone viral."

Under this reframing, we arrive at a major reconceptualization of the linguistic enterprise. Humboldt's "infinite use of finite means"—Berwick and Chomsky's "Basic Property"—is only part of the picture. It characterizes only the generative use of that subset of schemas that happen to be productive. In our picture, "lexical rules" are not relatively uninteresting add-ons to the all-important productive grammar. They embody all the partial regularities of the language, and even completely productive schemas can act as lexical rules. Hence, if anything is an add-on, it is the generative role of schemas. Only some schemas allow it. In short,

> **The generative property of language, the "infinite use of finite means," emerges from and rides on top of the system of lexical relations.**

This conclusion, which we might call the ***Relational Hypothesis***, is another of those "deeper generalizations" alluded to in Chomsky's codicil. It undermines the overwhelming emphasis in linguistic theory on productive phenomena. We therefore take it to be a primary responsibility of linguistic theory to characterize the patterns of relationships in the lexicon—which *include* all the patterns that also can appear in novel expressions.

To sum up, we are proposing a major realignment of linguistic theory. As we have stressed, many of these points are shared with other frameworks, especially Construction Grammar and Construction Morphology.

1. We have taken it as central that the lexicon is not simply a list of unstructured exceptions. Lexical items can have internal structure—morphological structure in the case of morphologically complex words, and syntactic structure in the case of idioms and other fixed expressions.
2. The lexicon includes a rich network of relational links among its items.
3. We have eliminated the strict distinction between lexicon and grammar, in that rules of grammar are in the lexicon.
4. Even within the lexicon, we have eliminated a hard distinction between words and rules, in that both are encoded in the same format, namely as pieces of linguistic structure connected by interface and relational links. The relevant distinction is whether a lexical item contains variables or not.
5. We have eliminated the distinction between "grammatical rules" and "lexical rules," in that both kinds of rules are in the lexicon, encoded in a common format. We have replaced this distinction with the distinction between the generative and relational functions of schemas.
6. We have weakened the morphology-syntax distinction, in that both are couched in terms of schemas, and both contain both productive and nonproductive schemas.
7. We have shown that the generative function of schemas is only half of the picture—or less than half—in that there are plenty of schemas that lack the generative function, especially in morphology. This dethrones the generativity of language from its dominating position in linguistic theory—although of course it is still an important issue. But lexical relations call for at least equal attention.

8. Beyond language

At the outset we advocated thinking of "knowledge of language" in terms of what is stored in long-term memory. Hence, we take Relational Morphology, complete with structured lexical entries, interface links, relational links, and schemas, to be a theory of one department of long-term memory. A natural question then is whether this theory scales up to the rest of language (as we have already suggested for syntax) and, beyond language, to other cognitive capacities.

A working hypothesis that we find attractive is that memory is memory is memory: pretty much the same organization is to be found in long-term memory for every cognitive domain. The differences among them lie in their specific primitives and affordances for combination, and in their interfaces to other domains. Language differs from other faculties, then, because it deals in syntactic, morphosyntactic, and phonological units, and because it interfaces with conceptual structure (meaning) on one end and with auditory

(or in the case of signed languages, visual) perception and motor control on the other—and because it is used for conventionalized communication.

This hypothesis invites us to seek similarities between linguistic memory and memory in other domains. Abstracting away from the content and function of linguistic memory to see its overall organization, we find general properties such as these:

1. A vast lexicon, with tens of thousands of items
2. Lexical items that involve multiple levels of representation, coordinated by interface links
3. Both free items (e.g., *pig*), which can occur independently, and bound items (e.g. *-ish*), which occur only attached to other material
4. Hierarchical constituent structure within items, on all three levels
5. Relational links among items that pick out shared structure
6. Regularities across items, which are picked out by schemas
7. Schemas that both motivate structure within stored items (their relational role) and assign structure to novel items (their generative role)

Very speculatively, let us see how many of these properties can plausibly be shared with other domains. Here, are some promising candidate domains.

8.1. Knowledge of music

1. Vast lexicon: Yes! One can recognize hundreds if not thousands of popular songs, folk songs, nursery rhymes, and, for some people, 45-minute symphonies and the like—to the extent that one can identify them immediately on hearing just a few random seconds of music (say, upon turning on the radio).
2. Lexical items that involve multiple levels of representation: Yes! Following Lerdahl and Jackendoff (1983), music cognition involves independent levels of grouping structure, metrical structure, and tonal hierarchy (or "prolongational reduction"), linked by a rich system of interface principles
3. Free versus bound items: It is hard to identify a musical counterpart of affixes. In Western tonal music, one possibility is the dominant harmony in a dominant-tonic cadence, whose function is to relax into the tonic. Another is the appoggiatura, a particular kind of fleeting dissonance that has to be resolved.
4. Hierarchical constituent structure within items: Yes! Music stored in memory has hierarchical structure on all three levels.
5. Relational links among items that pick out shared structure. Here, music might be a bit different from language: We might want to posit relational links *inside* a melody, picking out repetitions and variations of motives. For instance, one recognizes that the first two phrases of *Happy Birthday* are nearly identical. This sort of relational link does not appear in morphology or syntax, although it might occur in rhetorical speech registers with heightened affect such as poetry and preaching.

6. Regularities across items picked out by schemas: Yes! The well-formedness rules and preference rules of Lerdahl and Jackendoff can be recast as schemas. Schemas are also an attractive way to characterize conventionalized forms such as 32-bar popular song form, 12-bar blues, and classical sonata and minuet form.
7. Schemas both motivate structure within stored items and assign structure to novel items: Yes! There is no reason to believe that the structures of known music and newly experienced music are different in character. Well-known music may be represented more richly in memory, but according to the same principles.

8.2. Knowledge of objects

1. Vast lexicon: Yes! One can recognize thousands of objects and object types—without necessarily using linguistic labels.
2. Multiple linked levels of representation: Yes! One's knowledge of an object comprises not just how it looks, but it may also include how it feels (haptic representation), what sounds it makes, and how one uses it (action representations).
3. Free items and bound items: A stripe is bound: there cannot be a stripe without a surface. Holes, cracks, and dents are also candidates for bound items: there cannot be a hole without a volume in which it is situated. A handle may be physically free (it can be bought in the hardware store), but it is functionally bound (it has to be a handle of something, used to pick that something up).
4. Hierarchical constituent structure within items: Yes! Following Marr (1982), objects can be understood as having a hierarchical decomposition. A chair has legs, a seat, and a back; the back may have decorations; the decorations may be segmented into various parts, and so on. (See also Landau & Hoffman, 2012.)
5. Relational links among items that pick out shared structure: Yes? One can appreciate the similarity of structure between the seats of armchairs and wheelchairs, despite considerable superficial difference. One can appreciate the similarity in function between radically different kinds of bottle openers, lamps, or faucets.
6. Regularities among items picked out by schemas: Yes: Any sort of prototype representation (e.g., a Marr 3D model) is in effect a schema. Schemas can pick out generalizations about the layout of parts: for example, windows are normally placed in walls, not floors; walls of rooms normally meet at right angles. Rumelhart, 1980 uses the term *schema* in precisely this sense, speaking for instance of a schema for a face, with subschemas for noses and eyebrows and so on.
7. Schemas both motivate structure within stored items and assign structure to novel items: Again, there is no reason to think that when a novel object is committed to memory, the principles determining its structure should be any different.

8.3. Knowledge of geography and spatial layout

By this term we mean to encompass two different tasks that require similar sorts of knowledge: finding one's way from one place to another (Landau & Hoffman, 2012 again), and knowing where to find particular objects. (These might in fact be separate domains.)

1. Vast lexicon: Yes! How many places, streets, routes, does one know? How many associations of objects with places does one know? Where can the milk be found in the supermarket? Where does one keep the electrical tape? Where did I park this morning? (And possibly: In which publication did Lila Gleitman say such-and-such?) It is hard to know how to count, but we have huge amounts of such knowledge.
2. Multiple linked levels of representation: Hard to know.
3. Free items and bound items: Possible bound items include closets and suburbs, as appendices to a larger entity.
4. Hierarchical constituent structure within items: Yes. Where is the supermarket? Where is the milk within the supermarket? My brother's house is on a particular street in a particular neighborhood of a particular city, in a particular part of a particular state.
5. Relational links among items that pick out shared structure: Maybe. Similarities among particular airports or supermarkets?
6. Regularities among items, picked out by schemas: Yes. What one is likely to find in an airport versus in a bank versus in a restaurant?
7. Schemas both motivate structure within stored items and assign structure to novel items: Yes. One may use what one knows about supermarkets, based on supermarkets one has experienced, to make informed guesses about how a newly encountered supermarket is going to be organized, for instance that paper goods are likely to be grouped together.

8.4. Social knowledge

This again encompasses a number of issues: people one knows and what one knows about them; conventionalized social actions; and issues of moral value (Jackendoff, 2007).

1. Vast lexicon: Yes! One knows thousands of people to some degree or another, including not only how they look, but their personality and their social ties such as kin, spouses, ethnicity, religion, and occupation. One knows lots of rules of proper behavior, although it is hard to know how to count them: the fork goes to the left of the plate, people go to church on Sundays, you should give kids birthday presents . . .
2. Multiple linked levels of representation: Yes! A person is conceptualized as a linked physical instantiation (body) and a social presence (mind/spirit/soul) (Bloom,

2004; Jackendoff, 2007). Social actions have a physical instantiation linked to a social or moral value. For instance, shaking hands is a physical action fulfilling the social function of greeting or parting (among other things).

3. Free items and bound items: Moral values are not free-floating: They must be attached to some action.
4. Hierarchical constituent structure within items: Yes! Family structure, group structure (groups within groups within groups), authority (or rank) hierarchy.
5. Relational links among items picking out shared structure: Yes? Shared rank (all sergeants), shared occupation (all psycholinguists), shared intellectual ancestry (all former students of Lila Gleitman). Similarities among games (what soccer and hockey have in common; what pingpong and tennis have in common)
6. Regularities picked out by schemas: Yes! All customs, Minsky's (1975) frames (how a birthday party works), Goffman's (1974) frames (how a theater performance works), Rumelhart's (1980) schemas (how a buying/selling transaction works), and at a very general level, Fiske's (1991) four elementary form of human relations.
7. Schemas both motivate structure within stored items and assign structure to novel items: The stored items here would be particular occasions stored in episodic memory. They are presumably structured according to the same principles as one's assessment of ongoing events.

Our conclusion is that, at least at this impressionistic level, many general properties of linguistic knowledge are replicated in other domains of knowledge. Again, domains differ in what their structures are made of: phonological, morphological, and syntactic units in language; sequences of pitches with duration in music; visual/spatial units in object knowledge and geographical knowledge; persons, units of interpersonal behavior, and social/moral value in social knowledge. In order to develop this line of inquiry further, we need theories of representation in these other domains—theories that are comparable in sophistication with linguistic theory.

Although this last section has been speculative, it is a beginning step in testing a deeper generalization about the language faculty, moving toward integrating it with the rest of the mind. In turn, we take it that this reflects well on our linguistic theory: to the extent that the theory invites such integration, it encourages us to think we are on the right track.

Acknowledgments

We thank Jay Keyser, Peter Culicover, Barbara Landau, an anonymous reviewer, and members of the Moses Seminar at MIT for valuable comments on earlier versions of this work. RJ wishes to thank Peter Hagoort and the Max Planck Institute for Psycholinguistics for the opportunity to work on this project while a Visiting Fellow. JA is grateful to the Dutch national research organization NOW for a Veni grant, #275-70-036.

Notes

1. Other expositions of Relational Morphology, in various degrees of detail, appear in Jackendoff and Audring (2016), Culicover, Jackendoff, & Audring (2017), and our forthcoming book, *The Texture of the Lexicon*.
2. Jackendoff (2002) adds an interface directly between phonological and conceptual structure, to capture the relation between intonation contours in phonology and information structure in semantics.
3. The traditional notation, adopted in almost all frameworks, links the three components of this lexical item together by surrounding them with oversize square brackets. Our notation instead relies on the interface coindices to join the components together.
4. Although as stressed by Anderson (1992), it may be "atomic" from the point of view of phrasal syntax, which treats *piggish* just like any morphologically simple adjective such as *fat*.
5. A version of our relational links appears in the *network model* of Bybee (1995), although confined to phonology, and not linked directly to schemas, as ours are (see below).
6. Construction Grammarians would call (7) a ***construction***; some extend the term to words like (6) as well. To explicitly denote constructions with variables, we have appropriated the term *schema* from Construction Morphology (Booij, 2010), with which Relational Morphology shares a great deal. The term goes back in linguistics at least to Bybee & Slobin, 1982, who propose morphological schemas of approximately our sort, and it has a broader provenance in psychology that Rumelhart (1980) traces all the way back to Kant.
7. Our use of the terms "declarative" and "procedural" comes out of the computational tradition, and it differs from their use in the literature on memory. By "declarative," we intend the sort of mental representation in terms of which words are encoded—not just their meanings, which might be assigned to "semantic memory" but also their syntactic and phonological properties. By "procedural," we intend the sort of mental representation that encodes instructions for the active building of combinatorial structure. This may perhaps fall under the "procedural memory" of cognitive psychology and neuroscience.
8. Here, we diverge from the most popular versions of Construction Grammar, which require every construction to be associated with a meaning.
9. Such a switch in format can be seen for instance in Albright and Hayes's (2003) computational model of past tense acquisition, which otherwise is very much in tune with our approach.
10. It is sometimes suggested that idioms should be stored separately from words in a special "place" of their own. While this would allow the lexicon to consist only of words, it does not solve the problem of how idioms get their syntax and morphology. Moreover, idioms are intrinsically distinguished from words by virtue of

having internal syntactic structure, so they can be easily told apart without needing a special "place."

11. Granted, some of these words that we cite as nonexistent may be found in small quantities somewhere in some large corpus. But that does not mean they are necessarily stored in every individual's lexicon.
12. A different interpretation of Chomsky's proposal has emerged as Distributed Morphology (Halle & Marantz, 1993; Siddiqi, in press).
13. For differences between Unification and the Minimalist Program's Merge (Chomsky, 1995) as basic combinatorial functions in grammar, see Jackendoff (2011).
14. For instance, on November 1, 2016, in conversation with the first author, Jay Keyser coined the compound *theremin filibuster*. It was perfectly understandable in the context, (perhaps fortunately) now forgotten.
15. There are, however, a few idioms with other prepositions, such as *hand in hand* and *(tear) NP limb from limb*.
16. Other such cases include many of the "syntactic nuts" of Culicover (1999); some constructions discussed by Kay (2013); two constructions involving taboo terms, described in Hoeksema and Napoli (2008); and the place terms discussed by Culicover and Jackendoff (2005: 29).

References

Albright, A., & Hayes, B. (2003). Rules vs. analogy in english past tenses: A computational/experimental study. *Cognition*, *90*, 119–161.

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, *40*, 41–61.

Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology*, *3*, 232. https://doi.org/10.3389/fpsyg.2012.00232.

Anderson, S. (1992). *A-morphous morphology*. Cambridge, MA: Cambridge University Press.

Aronoff, M. (1976). *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Aronoff, M. (1994). *Morphology by itself*. Cambridge, MA: MIT Press.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*, 94–117.

Baayen, R. H., & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. *Linguistics*, *29*, 801–843.

Berwick, R., & Chomsky, N. (2016). *Why only us?*. Cambridge, MA: MIT Press.

Bloom, P. (2004). *Descartes' baby: How the science of child development explains what makes us human*. New York: Basic Books.

Booij, G. (2010). *Construction Morphology*. Oxford, UK: Oxford University Press.

Booij, G. (2017). Inheritance and motivation in Construction Morphology. In N. Gisborne & A. Hippisley (Eds.), *Defaults in morphological theory* (pp. 18–39). Oxford, UK: Oxford University Press.

Bresnan, J. (Ed.) (1982). *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.

Bresnan, J. (2001). *Lexical-Functional Syntax*. Malden, MA: Blackwell.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*, 425–455.

Bybee, J., & Slobin, D. (1982). Rules and schemas in the development and use of the English past tense. *Language*, *58*, 265–289.
Chomsky, N. (1957). *Syntactic structures*. The Hague, the Netherlands: Mouton.
Chomsky, N. (1959). Review of B. F. Skinner, *Verbal behavior*. *Language*, *35*, 26–58.
Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Chomsky, N. (1970) Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn. Reprinted in Chomsky, *Studies on semantics in generative grammar*, 11-61. The Hague, the Netherlands: Mouton, 1972.
Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, the Netherlands: Reidel.
Chomsky, N. (1986). *Knowledge of language*. New York: Praeger.
Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
Christiansen, M., & Arnon, I. (Eds.), (2017). Special issue: More than words: The role of multiword sequences in language learning and use. *Topics in Cognitive Science*, *9*(3).
Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, *22*, 991–1060.
Croft, W. (2001). *Radical Construction Grammar*. Oxford, UK: Oxford University Press.
Culicover, P. W. (1999). *Syntactic nuts*. Oxford, UK: Oxford University Press.
Culicover, P. W., & Jackendoff, R. (2005). *Simpler Syntax*. Oxford, UK: Oxford University Press.
Culicover, P. W., Jackendoff, R., & Audring, J. (2017). Multiword constructions in the grammar. *Topics in Cognitive Science*, *9*(1), 542–551. https://doi.org/10.1111/tops.12255.
Culicover, P. W., & Nowak, A. (2003). *Dynamical grammar*. Oxford, UK: Oxford University Press.
Fillmore, C., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language*, *64*, 501–538.
Fiske, A. (1991). *Structures of social life: The four elementary forms of human relations*. New York: Free Press.
Fodor, J. D. 1998. Learning to parse? In D. Swinney (Ed.), *Anniversary Issue of Journal of Psycholinguistic Research 27*, 285–318.
Gleitman, L., Gleitman, H., Miller, C., & Ostrin, R. (1996). Similar, and similar concepts. *Cognition*, *58*, 321–376.
Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Cambridge, MA: Harvard University Press.
Goldberg, A. (1995). *Constructions: A Construction Grammar approach to argument structure*. Chicago: Chicago University Press.
Goldberg, A. (2006). *Constructions at work*. New York: Oxford University Press.
Halle, M., & Marantz, A. (1993). Distributed morphology. In K. Hale & S. J. Keyser (Eds.), *The view from Building 20* (pp. 111–176). Cambridge, MA: MIT Press.
Haspelmath, M., & Sims, A. D. (2013). *Understanding morphology*. New York, NY: Routledge.
Hay, J., & Baayen, H. (2005). Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Science*, *9*, 342–348.
Hoeksema, J., & Napoli, D. J. (2008). Just for the hell of it: A comparison of two taboo-term constructions. *Journal of Linguistics*, *44*, 347–378.
Hoffman, T., & Trousdale, G. (2013). *The Oxford handbook of construction grammar*. Oxford, UK: Oxford University Press.
Humboldt, W. v. (1999 [1836]) *On the diversity of human language construction and its influence on the mental development of the human species* (orig. *Über die Verschiedenheit des menschlichen Sprachbaus und seinen Einfluss auf die geistige Entwicklung des Menschengeschlechts*). Cambridge, UK: Cambridge University Press.
Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language*, *51*, 639–671.
Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.
Jackendoff, R. (2002). *Foundations of language*. Oxford, UK: Oxford University Press.

Jackendoff, R. (2007). *Language, consciousness, culture*. Cambridge, MA: MIT Press.
Jackendoff, R. (2008). *Construction after construction* and its theoretical challenges. *Language*, *84*, 8–28.
Jackendoff, R. (2011). What is the human language faculty? Two views. *Language*, *87*, 586–624.
Jackendoff, R. (2015) In defense of theory. *Cognitive Science (special issue in commemoration of RJ's Rumelhart prize)*, *41*, Issue S2, 185–212. https://doi.org/10.1111/cogs.12324.
Jackendoff, R., & Audring, J. (2016). Morphological schemas: Theoretical and psycholinguistic issues. *The Mental Lexicon*, *11*, 467–493.
Jackendoff, R., & Audring, J. (in press) Relational morphology in the parallel architecture. In J. Audring & F. Masini (Eds.), *The Oxford handbook of morphological theory*. Oxford, UK: Oxford University Press.
Joshi, A. (1987). An introduction to tree-adjoining grammars. In A. Manaster-Ramer (Ed.), *Mathematics of language* (pp. 87–114). Amsterdam: Benjamins.
Kay, P. (2013). The limits of (Construction) grammar. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 32–48). New York: Oxford University Press.
Kiparsky, P. (1982) Lexical phonology and morphology. In I. S. Yang (Ed.), *Linguistics in the morning calm*, (pp. 3–91). Seoul: Hanshin.
Lakoff, G. (1970). *Irregularity in syntax*. New York: Holt, Rinehart and Winston.
Landau, B., & Gleitman, L. (2015). Height matters. In I. Toivonen, P. Csúri, & E. van der Zee (Eds.), *Structures in the mind* (pp. 187–210). Cambridge, MA: MIT Press.
Landau, B., & Hoffman, J. E. (2012). *Spatial representation: From gene to mind*. New York: Oxford University Press.
Langacker, R. (1987). *Foundations of Cognitive Grammar*. Vol. i. Stanford, CA: Stanford University Press.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
Levelt, W. J. M. (1989). *Speaking*. Cambridge, MA: MIT Press.
Marr, D. (1982). *Vision*. San Francisco: Freeman.
Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 425–432). Cambridge, MA: MIT Press.
Mohanan, K. P. (1986). *The theory of Lexical Phonology*. Dordrecht, the Netherlands: Reidel.
Moscoso del Prado Martin, F., Bertain, R., Häikiö, T., Schreuder, R., & Baayen, R. H. (2004) Morphological family size in a morphologically rich language: The case of Finnish compared with Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition 30*: 1271–1278.
Nooteboom, S., Weerman, F., & Wijnen, F. (Eds.) (2002). *Storage and computation in the language faculty*. Dordrecht, the Netherlands: Kluwer.
O'Donnell, T. (2015). *Productivity and reuse in language: A theory of linguistic computation and storage*. Cambridge, MA: MIT Press.
Pawley, A., & Syder, F. H. (1983). Two puzzles for lingusitic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–227). London: Longmans.
Pinker, S. (1999). *Words and rules*. New York: Norton.
Pollard, C., & Sag, I. (1994). *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
Pullum, G. (2007) The evolution of model-theoretic frameworks in linguistics. In J. Rogers & S. Kepser (ed.), *Proceedings of model-theoretic syntax at 10 (part of ESSLLI 2007)*, (pp. 1–10). Dublin: Trinity College Dublin.
Rumelhart, D. (1980). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Erlbaum.
Sadock, J. (1991). *Autolexical Syntax*. Chicago: University of Chicago Press.
Sapir, E. (1921). *Language: An introduction to the study of speech*. New York: Harcourt Brace.
Shieber, S. (1986). *An Introduction to unification-based approaches to grammar*. Stanford, CA: CSLI Publications.
Siddiqi, D. (in press) Distributed Morphology. In J. Audring & F. Masini (Eds.), *The Oxford handbook of morphological theory*. Oxford, UK: Oxford University Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, UK: Harvard University Press.

Van Valin, R., & LaPolla, R. (1997). *Syntax: Structure, meaning and function*. Cambridge, UK: Cambridge University Press.

Wasow, T. (1977). Transformations and the lexicon. In P. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal syntax* (pp. 327–360). New York: Academic Press.

Weinreich, U. (1969). Problems in the analysis of idioms. In J. Puhvel (ed.), *Substance and structure of language* (pp. 23–81). Berkeley: University of California Press. Reprinted in Weinreich, *On semantics*, 208–264. Philadelphia: University of Pennsylvania Press, 1980.

Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: MIT Press.