Relational Morphology in the Parallel Architecture

Ray Jackendoff and Jenny Audring

(to appear in J. Audring and F. Masini (eds.), *The Oxford Handbook of Morphological Theory*)

version of October 2016
* please check before citing *

ray.jackendoff@tufts.edu, j.audring@hum.leidenuniv.nl

1. Introduction

The Parallel Architecture (PA) (Jackendoff 1997, 2002, 2011) is a framework for understanding the organization of language and its place in the larger ecology of the human mind. It has met with considerable success in accounting for semantics, syntax and their interaction (Jackendoff 1983, 1990; Culicover and Jackendoff 2005). The present chapter sketches a morphological theory called Relational Morphology, based on the premises of the PA and drawing on the closely related approach of Construction Morphology (Booij 2010; Audring and Masini (this volume)). A fuller exposition appears in Jackendoff and Audring (forthcoming).[1]

A primary commitment of the PA is its thoroughgoing mentalism. The goal is to encode a speaker's knowledge of language in a fashion that not only accounts for linguistic structure, but that also bears meaningfully on psycholinguistic concerns such as the structure of memory, the processes of language comprehension and production, and the character of language acquisition. The theory aspires to be accountable to all the facts, and not to be limited by a competence/ performance or core/periphery distinction. Within morphology, this means that the theory must (a) encompass patterns of inflection, derivation, and compounding, from fully productive to incidental, (b) interact naturally with theories of syntax, semantics, and phonology, and (c) pay attention to issues in lexical processing and language acquisition.

PA interprets the term "knowledge of language" very literally – very psycholinguistically. "Knowledge" implies something stored in memory. From this perspective, the fundamental questions of linguistic theory can be formulated as:

- What linguistic units does one store in memory, and in what form?
- How are these units combined online to form novel utterances?
- How are these units acquired?

[1] We are delighted to thank Geert Booij and Jay Keyser for much discussion of this material and Peter Culicover for voluminous comments on an earlier version. RJ also thanks participants in the Tufts Linguistic Research Seminar (Ari Goldberg, Neil Cohn, Eva Wittenberg, Anita Peti-Stantić, Naomi Caselli, Anastasia Smirnova, Rabia Ergin, Irit Meir, Rob Truswell, and Katya Pertsova), who endured numerous barely formed versions of this material over a period of several years. We are especially grateful to Peter Hagoort and the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands for RJ's opportunity to spend time as a Visiting Fellow in the winters of 2014 and 2015, during which time we were able to develop much of the work reported here. JA is grateful to the Dutch national research organisation NWO for a Veni grant, #275-70-036.

Modern linguistics in the broad generative tradition has for the most part stressed the second of these questions, Humboldt's often-cited "infinite use of finite means" (1999 [1836]:91). In particular, many approaches to morphological theory have been couched in terms of how to build up morphologically complex words from smaller pieces. Nevertheless, it is hardly news that many morphologically complex words are semantically and/or phonologically idiosyncratic, making it necessary to store them (or parts of them) in some form or another. For instance, since the musical performance reading of *recital* cannot be derived from the meaning of verb *recite*, some version of it must be stored. Nevertheless, the relationship between *recite*, *recital*, and the affix *–al* is significant, and experimental results (e.g. Schreuder, Burani & Baayen 2003; Diependaele, Sandra & Grainger 2005) demonstrate that such relationships play a role in language processing. We conclude that morphological theory has to concern itself not just with the active generation of forms, but also with codifying the static relations among words and their constituents in memory. This is why we call our approach Relational Morphology (RM).

2. The place of morphology in the Parallel Architecture

A basic tenet of the Parallel Architecture is that linguistic structure is determined by three independent systems of representation – phonology, syntax, and semantics – and by the linkages among them. This is not a new idea: similar conceptions appear in Stratificational Grammar (Lamb 1966), Lexical-Functional Grammar (Bresnan 1982, 2001), Autolexical Syntax (Sadock 1991), Role and Reference Grammar (Van Valin and LaPolla 1997), and others. The upshot is an architecture like Figure 1. A well-formed sentence has well-formed structures in each of the three domains, plus well-formed links among the structures.

Phonological structures ⟵ Interface ⟶ Syntactic structures ⟵ Interface ⟶ Semantic structures

Figure 1. The Parallel Architecture

How do words fit into Figure 1? Within the Parallel Architecture, a word can be thought of as a small-scale interface rule, linking pieces of semantic, syntactic, and phonological structure, as in Fig. 2.

Phonology: /kæt/ ⟷ Syntax: N ⟷ Semantics: [CAT]

Figure 2. A word in the Parallel Architecture

Fig. 2 is a stereotypical lexical item: a word with structure in all three components, together making up a Saussurian sign. However, the lexicon also contains items that lack one or more of the levels. For instance, *hello*, *gosh*, and *oops* arguably have no syntactic category: they can serve as full utterances and they combine only paratactically. Other words, such as epenthetic *it*, complementizer *that*, and do-support *do*, are meaningless and serve only as grammatical "glue" (and hence are not Saussurean signs). Idioms such as *rock the boat* ('make trouble') are another sort of nonstereotypical lexical item: they have internal syntactic structure but noncompositional semantics (and hence within this idiom, *rock* and *boat* are not signs).[2]

Where is morphology in Figure 1? Traditional grammar treats morphology as a component of language distinct from phonology, syntax, and semantics. The Parallel Architecture suggests a different alignment. Instead of thinking of language as divided into four domains – phonology, morphology, syntax, and semantics – one should think of the architecture as a 3 x 2 matrix of components, as in Figure 3. Here, the grammar of words runs in parallel with the grammar of phrases, each involving phonological, syntactic, and semantic levels, with interfaces running between them. (Such an arrangement is prefigured in Bach 1983 and van der Hulst 2006.)

$$
\begin{tikzcd}
\text{Phrasal phonology} \arrow[r, leftrightarrow] \arrow[d, leftrightarrow] & \text{Phrasal syntax} \arrow[r, leftrightarrow] \arrow[d, leftrightarrow] & \text{Phrasal semantics} \arrow[d, leftrightarrow] \\
\text{Word phonology} \arrow[r, leftrightarrow] & \text{Morphosyntax} \arrow[r, leftrightarrow] & \text{Word semantics}
\end{tikzcd}
$$

Figure 3. Morphology in the Parallel Architecture

To understand these components and their relations, first consider morphosyntax – the internal syntactic structure of words. Derivational morphosyntax encodes the internal structure of complex words and the effects of word formation on syntactic category. Inflectional morphosyntax stipulates a language's dimensions of inflection, creating an n-dimensional array of syntactic categories and their associated inflectional features, each with a range of values, e.g. gender, number, and case for nouns; tense, person, and number for verbs.

In a syntactic structure, morphosyntax interfaces with phrasal syntax at the level of maximal X^0s, which constitute the largest morphological entities and the smallest syntactic ones. This level includes inflectional features, which are visible to both morphology and phrasal syntax. Aside from X^0 categories, morphosyntax and phrasal syntax do not share categories: affixes are found in morphosyntax but not phrasal syntax, while phrasal syntax has categories like NP, VP, and S, which with certain exceptions[3] play no role in morphosyntax.

Turning to the other components of Figure 3: word phonology concerns the phonological shape of words, including such matters as phonotactics, word stress, and vowel harmony. Phrasal phonology concerns phenomena such as phrasal stress and intonation contours. The two intersect at the level of the phonological word. However, the phonological phenomena of greatest interest

[2] Here PA diverges from the most popular versions of Construction Grammar (e.g. Goldberg 1996, 2006, Croft 2001, Boas and Sag 2012), in which every construction must be a sign, linking form (phonology and syntax) and function (semantics). In contrast, PA countenances meaning-free lexical items as well as meaningful ones. See Jackendoff 2013 for discussion.

[3] Exceptions include, for instance, the underlined parts of compounds such as *smoked pork shoulder boiled dinner*, *health and welfare fund*, and *"I have a dream" speech*. Selkirk 1982 has examples like *will-o'-the-wisp* and *ne'er-do-well*; Di Sciullo and Williams 1987 have French *trompe l'oeil*, *boit-sans-soif*, and English *matter-of-fact-ly* and *stick-to-it-ive-ness*. These can be treated as word-internal phrasal constituents, e.g. [$_{\text{Adv}}$ [$_{\text{NP}}$ *matter of fact*] [$_{\text{aff}}$ *ly*]], violating the canonical relation of morphosyntax and phrasal syntax. See Meibauer 2007 for some discussion.

to morphological theory are the interface principles that link word phonology to morphosyntactic structure. Morphosyntactic constituents stereotypically map one-to-one to phonological constituents, such that each piece of morphosyntax has a corresponding string of sounds. This matches a conventional item-and-arrangement conception of morphology. All the difficulties with IA morphology – and all the fun and danger – lie in noncanonical mappings between morphosyntactic features and phonology, in which the phonological form of inflected or derived words cannot be split cleanly into identifiable morphemes.

On the semantic side of Figure 3, similar considerations obtain. "Word semantics" specifies the possible semantic forms that words can take – the range of "lexical conceptual structures," in the sense of Jackendoff 1983.[4] What might be called "morphosemantics" is the interface mapping between morphosyntactic structure and word meanings. Morphological patterns can be used to express a heterogeneous collection of semantic functions such as causation, intention, time, aspect, evidentiality, and social formality, but also on occasion semantic factors farther afield (see Talmy 1978, Bauer 2002, and for compounding, Jackendoff 2010a). The morphosyntax-semantics interface is also responsible for the effects of morphological combination on argument structure. For example, event or process nominals such as *abandonment* preserve the argument structure of the corresponding verb, while agentive nominals like *baker* and result nominals like *inscription* denote one of the semantic arguments of the corresponding verb. Finally, word semantics canonically enters into phrasal semantics through the principle of compositionality. But there are also many noncanonical effects such as coercion, through which the meaning of a sentence can be more than the simple combination of its words, so-called "enriched composition" (Jackendoff 2002: 387-394, Audring & Booij 2016).

Overall, then, Relational Morphology sees the scope of morphology proper as encompassing morphosyntax plus its interfaces to phrasal syntax, word phonology, and word semantics. Of course, in order to understand an interface, one must understand both ends of what the interface connects. Hence the other components of the grammar cannot be neglected.

3. Productive and nonproductive schemas

3.1. Schemas vs. rules

An important tenet of PA and RM, shared with other theories such as LFG (Bresnan 2001), HPSG (Pollard and Sag 1994), Tree Adjoining Grammar (Joshi 1987), and Construction Grammar (including Construction Morphology), is that the grammar is stated entirely in terms of declarative patterns – which we call *schemas* – rather than in terms of procedural rules that apply in serial order to convert an "input" into an "output." All regularities in inflection and word-

---

[4] Lieber (this volume), citing among others Jackendoff 1990, characterizes lexical conceptual structure as the part of meaning relevant to syntax, a "semantic skeleton," contrasting with what she has called the "semantic body," "a less formal part that contains those encyclopedic elements of meaning that are of no syntactic relevance." One of the major arguments of Jackendoff 1983 (see also Jackendoff 2002, sections 9.6, 9.7) is that such a division is impossible. Rather, within the PA, lexical conceptual structure is to be understood as the entire meaning of words, whatever that proves to be. Only certain aspects of this are accessed by the syntax-semantics interface, for instance argument structure and animacy. These are the parts "relevant to syntax," but they are otherwise thoroughly integrated into the system of meaning, not a separate level or a separate kind of structure.

formation come to be stated in this fashion. For instance, the words in (1a) and (1b) are licensed by the schemas in (1c) and (1d).

(1) a. Semantics: [PLUR (CAT)]
Syntax: [$_{\text{Npl}}$ N, pl]
Phonology: /cats/

b. Semantics: [MAKE/BECOME (HARD)]
Syntax: [$_{\text{V}}$ A aff]
Phonology: /hardən/

c. Semantics: [PLUR (X)]
Syntax: [$_{\text{Npl}}$ N, pl]
Phonology: /...s/

d. Semantics: [MAKE/BECOME (X)]
Syntax: [$_{\text{V}}$ A aff]
Phonology: /...ən/

Schemas (1c) and (1d) have the same format as words, with the exception that they contain variables. Semantic variables are notated with capital letters, phonological variables are indicated by dots, and morphosyntactic variables are shown with their category label. Thanks to the shared format, schemas – like words – can be taken to be listed in the lexicon. Thus, as in Construction Grammar and Construction Morphology, both words and rules are stored pieces of linguistic structure, with complete continuity between them. A lexical item is more wordlike to the extent that its content is completely specific, such as the entry for *cats* in (1a); it is more rule-like to the degree that it contains variables, like the schema for the plural in (1c).

Schemas and procedural rules are sometimes thought to be notational variants. However, differences emerge as soon as one looks a little more deeply.

The first difference, of course, is that a theory based on procedural rules has two independent constructs – a lexicon and rules – whereas the constructional theory states words and schemas in a common format. As will be seen shortly, this is not just a difference in elegance.

A second difference concerns how stored lexical items are combined to produce novel complex words and utterances. In procedural theories, the rules build structures by applying in a determinate sequence, either from the top down or from the bottom up. In constructional theories, pieces of structure stored in the lexicon are “clipped together” by the process of **unification**, such that variables in schemas come to be instantiated. There is no inherent order of derivation: structures can be assembled from the bottom up, from the top down, or from left to right. This free order of assembly in constraint-based and constructional theories lends itself to being directly implemented in contemporary opportunistic theories of language processing (Jackendoff 2002, 2007; Sag 1992).

A third difference concerns the matter of storage. The usual interpretation of generative rules is that they are the source of combinatoriality in language: they produce novel composite forms, with complete generality. Yet there is abundant evidence that many composite items in language cannot be constructed by application of general rules. Examples include:

- Words with only partial compositional semantics. For example, *football* has something to do with feet and balls, but its full semantics is highly idiosyncratic (and different on different continents) and has to be learned and stored.
- Words with irregular phonological relations to their bases. One has to learn and somehow store the fact that the past tense of *sing* is *sang*, but the past tense of *cling* is *clung*.
- Words with predictable meanings, but whose existence has to be registered in the lexicon. For instance, deadjectival verbs in *–en*, e.g. *widen*, *brighten*, *redden*, are semantically uniform: '(cause to) become (more) A'. But not every predicted form exists, e.g. **louden*, **stouten*, **safen*.[5] So the existing forms have to be distinguished from the nonexistent ones.
- Words with an identifiable affix but a non-word base (so-called bound roots or cranberry morphs), such as *commotion* (**commote*) and *impetuous* (**impet*). These are far from rare; for instance, hundreds of English adjectives in *–ous* are of this character. They cannot be built by rule from smaller parts, since the base is not an independently existing part.

In addition, experimental results show that many composite forms, even regular ones, are stored (see, e.g., Giraudo & Grainger 2001; Andrews et al. 2004; Libben 2006; Baayen, Wurm & Aycock 2007; Kuperman et al. 2009 for compounds; and Baayen, Dijkstra and Schreuder 1997; Sereno and Jongman 1997; Baayen, McQueen, Dijkstra & Schreuder 2003; Sandra & Fayol 2003; Baayen 2007 for inflection). Consequently, a large number of complex lexical items must, for one reason or another, be stored in the lexicon.

Traditional procedural rules are unprepared for such a situation. On one hand, they overgenerate: any rule that produces *widen* will also produce **louden*, and any rule that produces *sang* will also produce **clang*. On the other hand, they undergenerate: they cannot predict the idiosyncratic semantics of *football*, and they have no base from which to derive *impetuous*. Moreover, in a procedural theory of grammar, relegating an item or pattern to the lexicon often means that it is no longer interesting. As Spencer (2013: 3) puts it: "Much of the derivational morphology discussed in the literature is ... of the occasional, accidental kind, and therefore of comparatively little interest to grammar writers (though it may be of interest to lexicographers, historians of language, psycholinguists, language teachers, and others)." The next section will show how a declarative theory based on schemas can offer a more inclusive model.

3.2. Lexical redundancy rules

A widely accepted approach to the difficulties enumerated above is the Lexicalist Hypothesis (Chomsky 1972): regular, predictable patterns belong to syntax, while more idiosyncratic

[5] This is true even taking into account the phonological constraints on the base: it must be monosyllabic and end with an obstruent.

patterns fall under the purview of "lexical redundancy rules" or simply "lexical rules," which apply before words are inserted into syntactic structures. This distinction became a foundational assumption of HPSG, LFG (where it was called Lexical Integrity), and many morphological theories (e.g. Aronoff 1976, Anderson 1992, Stump 1991, Spencer 2013). Depending on the theory, lexical rules can be viewed either as another layer of generative procedures, or else as establishing relations among stored items. The latter was the approach of Jackendoff 1975, which proposed an ancestral version of the schemas in, for instance, Bochner 1993, Booij 2010, and the present framework.

In RM, there is no occasion to posit separate lexical rules: all rules are 'lexical' in that they are pieces of linguistic structure containing variables, stored in the lexicon. Let us examine the relation between the adjective *hard* (2a), the verb *harden* from (2b) (=(1b)), and the schema (2c) (=(1d)) that captures the general relation that they instantiate. These are now enriched with a coindexation notation which allows us to specify which parts of each entry correspond to each other and to parts of other entries.

(2) a. Semantics: HARD_1 b. Semantics: $[\mathrm{MAKE/BECOME}\ (\mathrm{HARD}_1)]_2$
Syntax: A_1 Syntax: $[_{\mathrm{V}}\ \mathrm{A}_1\ \mathrm{aff}_3]_2$
Phonology: $/\mathrm{hard}/_1$ Phonology: $/\mathrm{hard}_1\ \mathrm{ən}_3/_2$

c. Semantics: $[\mathrm{MAKE/BECOME}\ (\mathrm{X}_x)]_y$
Syntax: $[_{\mathrm{V}}\ \mathrm{A}_x\ \mathrm{aff}_3]_y$
Phonology: $/..._x\ \mathrm{ən}_3/_y$

One should think of the coindices not as self-standing "lexical indices," but rather as marking the end of association lines: every subscript has to be paired with an identical subscript at the other end. In (2), subscript 1 connects the three components of *hard*. But it also connects the same three components inside the structure of *harden* – and it connects the corresponding parts of the two words. Subscript 2 connects the three components of the whole word *harden*. Subscript 3 connects the affix's phonology with its morphosyntax.

Turning to the schema in (2c), this says that English has verbs that (a) are based on an adjective with the meaning 'X', (b) have the affix /ən/ tacked onto the end of the base and (c) express the notion 'make or become (more) X'. The variable co-indices *x* and *y* allow this schema to be related to any word that has a corresponding pattern of co-indexation (e.g. *harden* and *widen*). On the other hand, the affix /ən/ itself is constant throughout all instances, so in schema (2c) as well as in *harden* (2b) it is encoded with a constant numerical co-index 3.

Schema (2c) is a generalization over existing words. As such, it expresses a relation among lexical items with corresponding structure. But it is not committed to productivity and it does not "generate" its instances. Hence, it functions as a lexical redundancy rule.

An important feature of this treatment is that schemas like (2c) do not require any additional components in the theory: they are formally identical to productive schemas. Consider the English regular plural schema in (3), adding the relevant coindexation to (1c) above.

(3) Semantics: $[\mathrm{PLUR}\ (X_v)]_z$
Syntax: $[_N\ N_v\ \mathrm{pl}_6]_z$
Phonology: $/..._v\ s_6/_z$

The schemas in (2c) and (3) are formally exactly parallel: they are pieces of phonological, syntactic, and semantic structure containing a variable on each level. The only major difference between them is that (3) is productive while (2c) is not. The next question then, is how the grammar makes this distinction.

3.3. Productivity

An easy solution is to mark each schema with a feature for whether it is productive or not. This is the approach taken in Booij 2010 (see also Audring & Masini this volume). A more fine-grained alternative is to mark productivity not on a schema as a whole, but on its variable, expressing the degree to which the variable is open to new lexical material (gradient productivity is assumed by, among others, Aronoff 1976, Lieber and Baayen 1991, and Baayen 1993; see also Hüning this volume and Bauer 2001 for discussion).

Marking productivity on the variable presents an interesting possibility: a schema could conceivably contain one productive (or open) variable and one nonproductive (or closed) variable. And indeed such schemas exist. For instance, English has four different patterns for naming geographical features, shown in (4a-d). The italicized words name the type of geographical feature.

(4) a. Arrowhead *Lake*, Loon *Mountain*, Wissahickon *Creek*, Laurel *Hill*, Sugar *Island*
b. *Mount* Everest, *Lake* Michigan, *Cape* Cod
c. the Indian *Ocean*, the Black *Sea*, the Hudson *River*, the White *Mountains*, the San Andreas *Fault*
d. the *Bay* of Fundy, the *Gulf* of St. Lawrence, the *Cape* of Good Hope

The choice of name is completely productive: if we wish to name a mountain for Morris Halle, we have no hesitation in calling it *Morris Mountain* or *Mount Morris*. On the other hand, the variable for the type of geographical feature is not productive. One has to learn which words go in which patterns; for instance **Mountain Morris* (pattern b) and **the Mountain Morris* (pattern c) are impossible. Hence the schemas for these patterns have one variable of each type, as in (5). Here the productive variable – the actual name – is notated with a double underline, and the nonproductive variable – the type of geographical feature – has a single underline; *the* and *of* in (5c,d) are constants.

(5) a. $\underline{\underline{N}}\ \underline{N}$ [e.g. *Loon Lake*]
b. $\underline{N}\ \underline{\underline{N}}$ [e.g. *Mount Washington*]
c. the $\underline{\underline{N}}\ \underline{N}$ [e.g. *the Lehigh River*]
d. the $\underline{N}$ of $\underline{\underline{N}}$ [e.g. *the Gulf of Mexico*]

Such a situation cannot be encoded if the distinction between productive and nonproductive is marked on the schema as a whole.

The patterns in (4)-(5) offer another kind of evidence for the PA's treatment of morphology and against a strict division between lexicon and grammar, as in the Lexicalist Hypothesis. On one hand, (4a,b) look like compounds. For example, they can be preceded by adjectival modifiers: *beautiful Arrowhead Lake*, *forbidding Mount Everest*. On the other hand, (4c,d) extend into the phrasal domain, because they have a determiner that can be followed by a modifying adjective: *the majestic Hudson River*, *the dangerous Bay of Fundy*. (4d) moreover has an *of*-phrase, characteristic of phrasal NP structure. The PA does not force us to form the first two "in the lexicon" and the other two "in the grammar," or to derive any of them from the others. Rather, all four of the schemas in (5) are in the lexicon: (5a,b) are morphosyntactic; (5c), in which the two nouns still form a compound, is a mixed morphosyntactic and phrasal schema; and (5d) is purely phrasal.

3.4 Two functions of schemas

The architecture laid out so far has three major advantages. First, having schemas and words in the same 'place' results in a network in which unproductive schemas can be linked to their instances without having to generate them. This allows for the relation to be partial, leaving room for idiosyncrasies such as the noncompositional semantics of *recital*, the irregular phonological relation of *destroy* and *destruction*, and the absence of a root as in *impetuous*. This solves the problems of undergeneration and overgeneration that we have found with traditional rules (section 3.1).

Second, productive schemas are also in the same place as their instances, so they, too, can be related to stored words. For instance, compounding is productive, and speakers encounter new instances all the time without notice. But at the same time there are thousands of conventionalized compounds like *football* with idiosyncratic meanings, which are linked to the compound schema. Similarly, the productive schema for plurals relates to lexically listed regular plurals such as *trousers* and *glasses* ('spectacles'). Even the morphologically regular *cats* has to be listed (though without its semantics) as part of the idiom *raining cats and dogs*, while still falling under the plural schema. Furthermore, experimental evidence suggests that highly frequent regular plurals are likely to be stored rather than generated online (Baayen 2007; Nooteboom, Weerman, and Wijnen 2002),[6] so that *cats* is more likely to be stored than, say, *coelacanths*. In each of these cases, the stored items fall under the very same schemas as items that are actively composed. The relational architecture provides a natural space for such phenomena.

A third advantage of the present architecture is that the distinction between productive and nonproductive schemas is not so stark. The major difference lies in the openness of the variable and, consequently, in the functions schemas can have. Schemas with an open variable have a **generative** function, used in creating new expressions – the function that has been most prominent in much linguistic theorizing. But in addition, all schemas – productive or

[6] More recently, psycholinguists see evidence for even more prolific storage. De Vaan et al. (2007) offer evidence that regular complex forms may already leave a trace in long-term memory after just a single exposure.

unproductive – have a **relational** function, in which they link to existing lexical items stored in the lexicon and capture generalizations among them. This leads to a notable change of perspective: one can think of productive schemas as just like nonproductive ones, except that in addition, they allow online instantiation of their variables. In other words, **productive schemas are "ordinary" schemas that have "gone viral."**[7] Such a view is possible only in a constructional framework like Construction Grammar/Morphology or the PA/RM. And this turns on its head the notion that the study of language should focus primarily on productive phenomena. Rather, the focus should be equally if not more on the relationships among items in the lexicon, that is, on the texture of linguistic knowledge.

3.5. Are nonproductive schemas necessary?

We take it as given that productive schemas are necessary in order to account for the construction of novel morphological forms. However, an important question for morphological theory is whether there actually are such things as <u>non</u>productive schemas. We have just pointed out some interesting consequences of assuming that they <u>are</u> necessary. However, this assumption is not universal. For instance, Pinker and colleagues (e.g. Pinker and Prince 1988, Pinker 1999), claim that there are rules (in our terms, schemas) for productive morphological patterns such as the English regular past tense, but that nonproductive patterns such as English ablaut past tenses are only a matter of association and analogy.

There certainly are cases of lexical relations that do not fall under a larger pattern and which might therefore just be matters of association. For instance, the pairs in (6) appear to be unique in their phonological relations to each other; there is no more general rule or schema that relates them.

(6) a. bomb/bombard
b. hate/hatred
c. laugh/laughter
d. humble/humility
e. bequeath/bequest
f. Glasgow/Glaswegian

A more questionable case is *horror/horrify/horrid/ horrible/horrific/horrendous*. The fact that these six words form a family might warrant a nonproductive schema for *horr-*, but the small size of the family might speak against it. (Marchand 1969: 5-6) shares our doubt.) On the other hand, schemas seem intuitively far more attractive for nonproductive patterns like [$_V$ A – *en*], with about 50 instances, and [$_N$ V – *ion*], with hundreds or thousands. Ultimately, we think, the question of whether there are schemas for particular cases is probably to be settled by psycholinguistics, not by morphological theory. We might even find that individuals differ in

[7] Could there also be schemas that have only the generative function and lack the relational function? No, because one can *always* store newly generated or encountered instances of a schema, which then fall under the schema's relational role.

how analytically they treat the language, and in how eager they are to form schemas (unconsciously, of course).[8]

One reason for positing nonproductive schemas comes from language acquisition. An essential part of this process is constructing (or discovering) the productive rules of the language, on the basis of primary linguistic input. Roughly, 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 (Tomasello 2003, Culicover and Nowak 2003). What is the form of such a hypothesis? Within the PA/RM and other constructionist theories, a hypothesis can be stated in the form of a tentative schema; its constants reflect the similarities among the words, and its variables reflect the differences among them.

However, learners have no way of knowing in advance whether a pattern they observe – and hence the hypothesized schema – will be fully productive or not. So they will inevitably create a lot of schemas that fail the criteria for productivity (whatever these criteria may be – see, for instance, Baayen 1993, Yang 2005, O'Donnell 2015). What happens to failed schemas? If the brain does not countenance nonproductive schemas, these hypotheses have to be expunged. But we see little reason for the brain to throw out information about linguistic patterns if it can be useful. (For more discussion, see the next section and Jackendoff and Audring forthcoming.)

In RM's view of schemas, productive and nonproductive schemas are in exactly the same format. If a learner extracts a pattern as a 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 Rule and No Rule, as in Pinker's approach. It is just a matter of determining the proper diacritic on the schema's variable – open vs. closed. 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 have to relinquish its status as a lexical redundancy rule. Rather, as suggested in section 3.3, it retains the function of capturing generalizations among stored items.

One way that nonproductive schemas can be useful is 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.

[8] For an experimental study of individual variation in the systematicity of compounding, see Gleitman and Gleitman 1970. Pinker 1999 discusses differences in storage versus computation of English past tenses across different neurological populations.

4. Formalizing lexical relations

4.1. Inheritance with impoverished entries and full entries

As indicated by the name, Relational Morphology puts a strong emphasis on the relations between lexical items. We now begin to dig a little deeper into the meaning of our formalization.

A common position (e.g. Booij 2010) is that complex words, their bases and their schemas are related through an **inheritance hierarchy**, as illustrated in (7). The items lower in the hierarchy are taken to inherit from items higher in the hierarchy to which they are connected. Thus *harden* inherits from *hard*, *whiten* inherits from *white*, and both inherit from the schema.

(7) $_{A}$hard [$_{V}$ A-en] $_{A}$white

[$_{V}$ $_{A}$hard en] [$_{V}$ $_{A}$white en]

Inheritance is especially attractive because it has also been frequently invoked in the organization of concepts. Hence it is a domain-general theoretical construct that requires no special machinery for language alone (Murphy 2002). But what does inheritance mean? What do the lines signify? We will explore two interpretations of inheritance, rejecting one and developing a more flexible version of the other.

A common interpretation of inheritance (e.g. Collins and Quillian 1969, Pollard and Sag 1994, and in a sense Chomsky and Halle 1968; see also Brown (this volume)) is that the lexicon is maximally economical: any information present in a higher node of the hierarchy does not have to be specified in a lower node. For instance, *harden* inherits everything from the two higher nodes and therefore can be listed something like this:

(8) $_{A}$hard [$_{V}$ A-en]

[. .]

The base of a word like *impetuous* is not listed in the lexicon, so it inherits only from the schema for the affix and has to list the rest, as in (9).

(9) [$_{A}$ N – ous]

[impet .]

Following Jackendoff 1975, we call this the **impoverished entry theory**: the idea is that lexical items contain only information that cannot be inherited from elsewhere.

Despite its intuitive appeal, there are numerous reasons to reject this position. We will mention two (others appear in Jackendoff 1975 and Jackendoff and Audring forthcoming). First, it implies that in order to retrieve a complex item such as *harden* from long-term memory, one must fill in its missing content from its superordinates in the hierarchy. Hence a complex item should always take longer to process than its base. However, experimental evidence proves this prediction false. For example, Baayen, Dijkstra and Schreuder 1997 report that if a Dutch noun stem (appearing in either singular, plural, or diminutive) is high frequency, and if in addition the plural is more frequent than the singular, then the plural is accessed as quickly as the singular – not more slowly, as the impoverished entry theory would predict.

Second, consider the theory of acquisition. In order to construct a new schema, one must generalize over existing lexical items whose details are present in memory. The impoverished entry theory has to claim that once the schema is established, the redundant details of all the instances used to establish it are erased from memory, in order to optimize the lexicon. We find this implausible (though psycholinguistic evidence might prove us wrong). Similarly, in the course of acquiring a new complex word, one must first discover its details, and only then determine what schemas it falls under and what its base is. The impoverished entry theory entails that as soon as one establishes the new word's relation to a base and to one or more schemas, all the redundant features are immediately expunged. Again, we find this implausible. (Similar arguments can be found in Langacker 1987 and Booij to appear.)

The basic difficulty with the impoverished entry theory is that it assumes there to be a premium on economy and elegance, sometimes formalized as "minimum description length." Our question is whether economy is the right criterion when it comes to storage in the brain. A plausible alternative is that the brain embraces redundancy, at least up to a point. For instance, languages seem to have no problem marking thematic roles redundantly, through word order, case marking, and verb agreement.

Given these precedents, we find it attractive to consider a **full-entry** theory, in which lexical items are encoded in their entirety, even where redundant.[9] For instance, the full-entry theory says that the lexical entry of *harden* is like (2b) (repeated in (10b) below), with its full structure, rather than like (8), in which it has been evacuated of content. The role of a schema, then, is not to permit material to be omitted from other entries, but to confirm or codify or **motivate** generalizations among lexical entries. The notion of motivation is invoked widely in morphology, e.g. Booij 2010, to appear, Cognitive Grammar (Lakoff 1987, Goldberg 1995, Radden and Panther 2004), and indeed by Saussure (1915, 133).

Relational Morphology fleshes out the notion of motivation in terms of shared structure. To illustrate, we return to our treatment of *harden*, repeated here.

(10)	a.	Semantics:	HARD_1	b. Semantics:	$[\text{MAKE/BECOME} (\text{HARD}_1)]_2$
		Syntax:	A_1	Syntax:	$[_V\ A_1\ \text{aff}_3\ ]_2$

[9] We acknowledge that there are important questions about what it means to code an item "in its entirety." Does that include phonetic detail? Does it include semantic detail that might be termed "real-world knowledge"? We must leave these questions open. However, we do not subscribe to an exemplar-theoretic view in which one stores all experienced tokens in full detail and makes no abstractions.

Phonology: /hard/$_1$ Phonology: /hard$_1$ ən$_3$ /$_2$

c. Semantics: [MAKE/BECOME ($\underline{X}_x$)]$_y$
Syntax: [$_V$ $\underline{A}_x$ aff$_3$]$_y$
Phonology: / $\underline{\ldots}_x$ ən$_3$ /$_y$

The shared structure is encoded in the coindices. Subscript 1 ties together the semantics, syntax, and phonology of the word *hard* in (10a). It also ties together parallel constituents of *harden* in (10b). At the same time, though, it links these constituents of (10b) to the corresponding parts of (10a), thereby identifying the structure shared between the two entries. Similarly, subscript 3 picks out structure shared between *harden* and schema (10c): the affix and its phonological realization. Finally, as suggested above, the variable subscripts *x* and *y* in (10c) indicate structure shared with anything that has parallel structure – i.e. all the instances of the schema. An immediate advantage of the notation in (10) over standard notations for inheritance such as (7) is that it enables us to pinpoint the regions of similarity among items.

4.2. Inheritance without inherent directionality

However, inheritance as usually conceived – even within full-entry approaches – is not general enough for the full range of morphological relations. First, inheritance is usually considered to be asymmetrical and top-down, such that the general fills in the particular: a complex item such as *harden* inherits from its base (*hard*) and the schema [$_V$ A – *en*]. However, as we observed above, schemas have to be acquired from the bottom up, by generalizing over some collection of previously stored examples. Hence in a sense, the instances motivate the schema, rather than (or in addition to) the other way around. This conundrum does not arise with the coindexing notation in (10), which is not inherently directional. We do not have to decide whether (10c) is motivated by (10b) or vice versa: the schema and its instances mutually motivate each other.

A further problem occurs with pairs like *assassin/assassinate*. Phonologically, *assassinate* is clearly built on the base *assassin* and should inherit from it. But semantically the relation goes the other way: an assassin is someone who assassinates people. If the phonological relation actually mirrored the semantics, we would get an ordinary agentive nominal **assassinator*. Similarly for *linguist/linguistics*: in the standard version of inheritance, *linguist* would be the ancestor of *linguistics* on the phonological level, but its descendant on the semantic level.

(11) shows how such cases can be treated in the coindexing notation. The semantics of (11b), coindexed 8, is shared with part of (11a), while the phonology of (11a), coindexed 7, is shared with part of (11b). (Coindex 9, which identifies the affix, is shared with the *–ate* schema, not shown here.) Beyond this, there is no need to say which word is derived from which.

(11) a. Semantics: [PERSON WHO [MURDERS POLITICIAN]$_8$]$_7$
Syntax: N$_7$
Phonology: /əsæsən/$_7$

b. Semantics: [MURDER POLITICIAN]$_8$

Syntax: $[_V\ N_7 - \text{aff}_9\ ]_8$
Phonology: / //əsæsən/$_7$ /eyt/$_9$ /$_8$

In these sorts of examples, the PA notation is invaluable, because it makes it possible to correlate phonology, syntax, and semantics independently.

4.3. A-morphousness

Next consider a word like *gorgeous*, which has a legitimate affix attached to a non-word base. Like all words with the affix *–ous*, this word is an adjective. But since the base is not a word, it has no syntactic category. There are hundreds of such *–ous* words (e.g. *scrumptious*, *curious*, *impetuous*, *supercilious*, *meticulous*), in addition to those with a genuine noun base, such as *joyous*. In RM, we can encode *gorgeous* as (12a), where the morphosyntax of the base is blank: this part of the word is in effect unparsed in morphosyntax, although it does have phonology. (12b) offers a schema that encompasses both lexical and nonlexical bases; the angle brackets < > indicate optional specifications on the variable.

(12) a. Semantics: BEAUTIFUL_{10}
Syntax: $[_A - \text{aff}_{11}\ ]_{10}$
Phonology: /gordʒ əs$_{11}$ /$_{10}$

b. $[_{\text{Property}}$ <PERTAINING TO X_z>$]_w$
$[_A < N_z > \text{aff}_{11}\ ]_w$
/ …$_{<z>}$ əs$_{11}$ /$_w$

Since *gorge-* is not represented in morphosyntax, one might consider *gorgeous* to be partially "a-morphous" in the sense of Anderson 1992.

The words in (13) present a more extreme case.

(13) million, billion, trillion, …; zillion, godzillion, kajillion

While the pattern is easily recognizable and can be creatively extended, as in the last three examples, the morphological status of *–illion* is questionable. It cannot be a suffix, as that would imply that *m-*, *b-*, *z-* and so on are bound roots, many of which fail to meet the phonological requirements for a root. The alternative, treating *–illion* as a root and the consonants as prefixes, would imply a cluster of nonce prefixes that do not occur anywhere else in the language. Instead of settling for either of these unsatisfactory analyses, we can say that these particular words are completely a-morphous. The phonological string *–illion* is perhaps associated with a meaning 'very large number,' but there is no morphosyntactic category such as Affix or Noun associated with this meaning. (14a) shows an entry for *trillion* and (14b) is a possible *–illion* schema.

(14) a. Semantics: $10^{12}{}_{12}$
Syntax: Numeral_{12}
Phonology: / /tr/ /illion/$_{13}$ /$_{12}$

b. Semantics: $\underline{\text{LARGE NUMBER}}_x$
Syntax: $\underline{\text{Numeral}}_x$
Phonology: /$\underline{\ldots}$/ /illion/$_{13}$ /$_x$

(14) treats *trillion* as simply a morphosyntactic Numeral with no internal morphosyntactic structure. At the same time, the phonology and semantics are correlated, and the schema is available for coining jocular number words such as *kajillion*. So here is a fully a-morphous

example in Anderson's sense. A similar solution can be applied to phonaesthemic patterns such as the *gl*- in *glimmer*, *glow*, *gleam*, and so on.

Anderson (1992) advocates the position that morphologically complex items have no internal structure. We are urging a more nuanced view: items like *harden* and *joyous* do have internal morphosyntactic structure, but items like *gorgeous* have only partial morphosyntactic structure, and items like *trillion* lack it altogether. Nevertheless, all of them can be partially motivated by schemas. This heterogeneous view of morphosyntactic complexity is of a piece with RM's overall outlook on the character of the lexicon.

4.4. Sister relations

The non-word bases in the previous section bring us to another case of motivation that poses a serious problem for inheritance. Consider pairs like *ambition/ambitious* and *cognition/cognitive*, which are clearly related, but for which there are no lexical bases **ambi(t)* and **cogni(t)* that both members of the pair can inherit from. The standard construal of inheritance requires us to say that one of the pair is basic and the other "derived." Yet there is no clear way to say whether *ambition* inherits from *ambitious* or vice versa. We would prefer simply to say that they share structure and are related as equal "sisters," without a hypothetical bound root as the "mother" that ties them together.

The coindexing notation makes it straightforward to express such sister relations, for instance as in (15). The sisters share part of their semantics (subscript 14) and part of their phonology (subscript 18). But they do not share their affixes: subscript 16 is coindexed with the *–tion* schema (not shown) and subscript 17 with the *–tious* schema (an allomorph of the *–ous* schema in (12b)). Finally, since **ambi* is not a word and has no part of speech, we leave its morphosyntax blank, as we did for *gorgeous* in (12a). Crucially, unlike *hard*/*harden* (10a,b), neither entry is contained in the other, and again we do not need to decide which is derived from the other. This is what makes them sisters rather than a standard mother-daughter inheritance configuration.

(15)	a.	Semantics:	DESIRE$_{14}$	b. Semantics:	[HAVING (DESIRE$_{14}$)]$_{15}$
		Syntax:	[$_{N}$ – aff$_{16}$]$_{14}$	Syntax:	[$_{A}$ – aff$_{17}$]$_{15}$
		Phonology:	/ /æmbɪ/$_{18}$ /ʃən/$_{16}$ /$_{14}$	Phonology:	/ /æmbɪ/$_{18}$ /ʃəs/$_{17}$ /$_{15}$

4.5. Sister relations among schemas

As mentioned in section 3.4, some sister relations, such as *bomb/bombard*, appear to be one-off pairings. Others are more systematic. A case discussed by Booij 2010 is the pairing of names of ideologies with names of their adherents, such as in (16):

(16)
a. pacifism/pacifist
b. altruism/altruist
c. solipsism/solipsist
d. impressionism/impressionist

To express this generalization, we have to establish a relation not only between two sister words as in (15), but also between two sister schemas: the *–ism* schema for ideologies and the *–ist* schema for their adherents. Booij (2010: 31-36, Booij & Masini 2015) calls this configuration a "second-order schema"; we call it **sister schemas** (see also Nesset 2008, Kapatsinski 2013, and Audring & Masini this volume). While the idea has been developed elsewhere, RM offers a more precise way to notate the relations.

(17) a.	Semantics:	$\underline{\text{IDEOLOGY}}_x$	b. Semantics:	[ADHERENT ($\underline{\text{IDEOLOGY}}_x$)]$_z$
	Syntax:	$[_N - \text{aff}_{19}\ ]_x$	Syntax:	$[_N - \text{aff}_{20}\ ]_z$
	Phonology:	/ $\underline{/\ldots/}_y$ /ɪzm̩/$_{19}$ /$_x$	Phonology:	/ $\underline{/\ldots/}_y$ /ɪst/$_{20}$ /$_z$

Let us unpack this. (17a) says there can be nouns that denote an ideology and that end with the affix *–ism* (coindex 19). The variable coindex x ties the three components together. (17b) says that there can be nouns that denote the adherents of an ideology and that end with the affix *–ist* (coindex 20); the variable coindex z ties these three components together.

So far these schemas are just like the *–ous* schema in (12b). But what makes them more interesting is that IDEOLOGY in (17a) is linked to IDEOLOGY in (17b) by the variable coindex x, and the part of the phonology in (17a) that precedes the affix is linked to the corresponding part in (17b) by the variable coindex y. Thus this pair of schemas together says that for any noun ending in *–ism* that denotes an ideology, it should not be surprising to find another noun ending in *–ist* that is phonologically the same up to the affix (coindex y), and that denotes an adherent of that very ideology (coindex x) – and, since coindexation is nondirectional, vice versa as well.

Booij notates sister schemas as a special relation between schemas:
< Schema A > ≈ < Schema B > . In the present approach, sister schemas are a generalization of sister relations between words. Where the sister relation in (15) links numerical coindices across two words, the second-order schema in (17) links variable co-indices across two schemas. Instances of the two schemas can be connected as sisters as well. Thus the present notation achieves a somewhat greater degree of generality, and further brings out the continuity between words and rules.

5. Summary and conclusions

We have explored here the outlook of Relational Morphology, a theory of morphology grounded in the framework of the Parallel Architecture. The fundamental goal is to describe what a language user stores in memory and in what form, and to describe how this knowledge is put to use in constructing and comprehending novel utterances. A basic tenet of RM, following PA, is that knowledge of language is segregated into phonological, syntactic, and semantic/conceptual structures, plus interfaces between them that enable sound, morphosyntactic structure, and meaning to be related to each other. Words function as small-scale interface rules, establishing links among pieces of structure in the three domains.

Within this outlook, morphology emerges as the grammar of word-size pieces of structure and their constituents, comprising morphosyntax and its interfaces to word phonology, lexical semantics, and phrasal syntax. Canonical morphology effects a straightforward mapping between these components; irregular morphology is predominantly a matter of non-canonical mapping between constituents of morphosyntax and phonology.

As in Construction Morphology, RM encodes rules of grammar as schemas: pieces of linguistic structure containing variables, but otherwise in the same format as words – that is, the grammar is part of the lexicon. Hence there is no principled distinction between the formalisms for words and for rules, aside from the presence or absence of variables – a simplification of the repertoire of theoretical constructs.

Productive schemas serve two functions. In their **generative** function, they are used to build novel utterances by "clipping" pieces of structure together, one piece instantiating a variable in the other through unification. In their **relational** function, they serve to motivate relations among items stored in the lexicon. Morphological patterns that are *not* productive can also be described in terms of schemas. Such schemas are formally parallel to those for productive patterns, except that they have only the relational function. We have argued that there is no principled distinction between these two sorts of schemas, aside from a diacritic on their variables expressing their degree of openness – again a simplification of the theoretical apparatus. Because *all* schemas participate in the relational function, we conclude that morphological theory should focus on expressing lexical relations at least as much as on the online construction of novel forms. Such a focus reveals the lexicon to be richly textured, not the unstructured list that many linguists have made it out to be.

Finally, we have addressed how lexical relations are to be expressed. Beginning with the well-known mechanism of inheritance, we have shown that an impoverished-entry theory of inheritance is inappropriate for a variety of reasons, and that a full-entry theory is more satisfactory. However, the full-entry theory itself is not general enough: it is unable to deal with lexical relations that are nondirectional, multidirectional, or symmetrical. We have proposed to express motivation (a generalized form of inheritance) in terms of shared structure, and we have introduced a notation that enables us to flexibly pinpoint the regions of commonality between pairs of words, between words and schemas, and between pairs of schemas.

The challenge for this theory is to apply it to the full range of issues investigated in current morphological theorizing. Many representative phenomena are addressed in Jackendoff and Audring (forthcoming), for instance zero morphology, stem allomorphy, blends, truncations, and inflectional classes. However, even in the present chapter, we have been able to touch on some telling phenomena that have not been prominent in the literature, using a minimum of theoretical machinery. More broadly, we believe that the way the issues have been couched here, emphasizing the theory's consistency with the larger framework of the Parallel Architecture, can help build bridges between linguistic theory, psycholinguistics, and cognitive neuroscience.

References

Anderson, Stephen. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, 16, 285–311.

Aronoff, Mark. (1976). *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Audring, Jenny, and Francesca Masini. (this volume, forthcoming) Construction Morphology. In Jenny Audring and Francesca Masini (eds.) *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press.

Baayen, R. H. (1993). On frequency, transparency and productivity. In: Booij, G., van Marle, J. (eds.), *Yearbook of Morphology 1992*. Dordrecht : Kluwer, 181-208.

Baayen, R. H. (2007). Storage and computation in the mental lexicon. In G. Jarema and G. Libben (eds.), *The Mental Lexicon: Core Perspectives*, Elsevier, 81-104.

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*, 36, 94-117.

Baayen, R. H., McQueen, J., Dijkstra, T., and Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In Baayen, R. H. and Schreuder, R. (eds.) *Morphological structure in language processing*. Mouton de Gruyter, Berlin. 355-390.

Baayen, R. H., Wurm, L. H., & Aycock, J. (2007). Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities. *Mental Lexicon*, 2, 419-463.

Bach, Emmon. 1983. On the relationship between word-grammar and phrase-grammar. *Natural Language and Linguistic Theory* 1.65-89.

Bauer, Laurie. (2001). *Morphological Productivity*. Cambridge: Cambridge University Press.

Bauer, Laurie. (2002). What You Can Do with Derivational Morphology. In Sabrina Bendjaballah, Wolfgang U. Dressler, Oskar E. Pfeiffer, & Maria D. Voeikova (eds.) Morphology 2000: Selected papers from the 9th Morphology Meeting, Vienna, 24–28 February 2000.

Boas, Hans, and Ivan Sag (eds.). (2012). *Sign-Based Construction Grammar*. Stanford: CSLI.

Bochner, Harry. (1993). *Simplicity in Generative Morphology*. Berlin and New York: Mouton de Gruyter.

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

Booij, Geert (to appear), Inheritance and motivation in Construction Morphology. In Nikolas Gisborne and Andrew Hippisley (eds.), Default inheritance. Oxford: Oxford University Press.

Booij, Geert, and Francesca Masini. (2015) The role of second order schemas in the construction of complex words. In Laurie Bauer, Livia Kőrtvélyessy and Pavol Štekauer (eds.) *Semantics of complex words.* Cham etc: Springer, 47-66

Bresnan, Joan (ed.). (1982). *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

Bresnan, Joan. (2001). *Lexical Functional Syntax*. Oxford: Blackwell.

Brown (this volume, forthcoming) Network Morphology. In Jenny Audring and Francesca Masini (eds.) *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press.

Chomsky, Noam. (1972). *Studies on Semantics in Generative Grammar*. The Hague: Mouton.

Chomsky, Noam, and Morris Halle. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Collins, Allan M., and M. Ross Quillian. (1969). Retrieval time from semantic memory. *Journal of Verb Learning and Verbal Behavior* 8: 240-247.

Croft, William. (2001). *Radical Construction Grammar*. Oxford: Oxford University Press.

Culicover, Peter, and Ray Jackendoff. (2005). *Simpler Syntax*. Oxford: Oxford University Press.

Culicover, Peter, and Andrzej Nowak. (2003) *Dynamical Grammar*. Oxford: Oxford University Press.

de Vaan, L., Schreuder, R. and Baayen, R. H. (2007). Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon* 2, 1-23.

Diependaele, Kevin; Sandra, Dominiek; Grainger, Jonathan. (2005). Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes*, Vol 20(1), 75-114.
Di Sciullo, Anna Maria, and Edwin Williams. (1987). *On the Definition of Word*. Cambridge, MA: MIT Press.
Giraudo, H. & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, Vol. 15(4/5), 421-444.
Gleitman, Lila, and Henry Gleitman. (1970). *Phrase and Paraphrase: Some Innovative Uses of Language*. New York: W. W. Norton.
Goldberg, Adele. (1996). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
Goldberg, Adele. (2006). *Constructions at Work*. Oxford: Oxford University Press.
Hüning, M. (this vol, forthcoming) Word-formation theory and diachronic change. In Jenny Audring and Francesca Masini (eds.) *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press.
Humboldt, Wilhelm von. (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*). Michael Losonsky (Ed.), Cambridge: Cambridge University Press
Jackendoff, Ray. (1975). Morphological and semantic regularities in the lexicon. *Language* 51, 639-71. Reprinted in Jackendoff 2010b, 40-84.
Jackendoff, Ray. (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.
Jackendoff, Ray. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
Jackendoff, Ray. (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
Jackendoff, Ray. (2002). *Foundations of Language*. Oxford: Oxford University Press.
Jackendoff, Ray. (2007). A Parallel Architecture Perspective on Language Processing, *Brain Research* 1146, 2-22.
Jackendoff, Ray. (2010a) The ecology of English noun-noun compounds. In Jackendoff 2010b, 413-45.
Jackendoff, Ray. (2010b). *Meaning and the Lexicon: The Parallel Architecture 1975-2010*. Oxford: Oxford University Press.
Jackendoff, Ray. (2011). What is the human language faculty? Two views. *Language* 87, 586-624.
Jackendoff, Ray. (2013). Constructions in the Parallel Architecture. In T. Hoffmann and G. Trousdale (eds.), *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press, 70-92.
Jackendoff, Ray, and Jenny Audring (forthcoming). *The Texture of the Mental Lexicon: Morphology in the Parallel Architecture*.
Joshi, Aravind K. (1987) An introduction to Tree-Adjoining Grammars. In A. Manaster-Ramer (ed.), *Mathematics of Language*. Amsterdam: Benjamins, 87-114.
Kapatsinski, Vsevolod. (2013). Conspiring to mean: Experimental and computational evidence for a usage-based harmonic approach to morphophonology. *Language* 89(1): 110-148.
Kuperman, V., Schreuder, R., Bertram, R., Baayen, H.R. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception & Performance*, 35(3), 876-895.
Lakoff, George. (1987). *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press.
Lamb, Sydney. (1966). *Outline of Stratificational Grammar*. Washington: Georgetown University Press.
Langacker, Ronald. (1987). *Foundations of Cognitive Grammar*, vol. i. Stanford: Stanford University Press.
Libben, G. (2006). Why study compound processing? In G. Libben & G. Jarema (eds.) *The representation and processing of compound words*. Oxford: Oxford University Press. 1-22.
Lieber, Rochelle. (this volume). Theoretical issues in morphology: Word formation.
Lieber, R., and H. R. Baayen. (1991). Productivity and English derivation: a corpus-based study. *Linguistics* 29: 801-844.

Baayen, H. & R. Lieber (1991). Productivity and English derivation: a corpus-based study. *Linguistics* 29: 801-844.
Marchand, H. (1969). *The categories and types ofpresent-day English word-formation.* 2nd edn. Munich: Bech.
Murphy, Gregory. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.
Meibauer, Jörg. (2007). How marginal are phrasal compounds? Generalized insertion, expressivity, and I/Q-interaction. *Morphology* 17: 233–259.
Nesset, Tore. (2008). *Abstract phonology in a concrete model: Cognitive linguistics and the morphology-phonology interface*. Berlin / New York: Mouton de Gruyter.
Nooteboom, Sieb, F. Weerman, and F. Wijnen (eds.). (2002). *Storage and Computation in the Language Faculty*. Dordrecht: Kluwer.
O'Donnell, Timothy. (2015). *Productivity and Reuse in Language: A Theory of Linguistic Computation and Storage.* Cambridge, MA: MIT Press.
Pinker, Steven. (1999). *Words and Rules*. New York: Basic Books.
Pinker, Steven, and Alan Prince. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition*.* *Cognition* 26, 195-267.
Pollard, Carl, and Ivan Sag. (1994). *Head-Driven Phrase Structure Grammar*, Chicago: University of Chicago Press.
Radden, Günther and Klaus-Uwe Panther (eds.) (2004). Studies in Linguistic Motivation. Berlin/ New York: Mouton de Gruyter.
Sadock, Jerrold. (1991). *Autolexical Syntax*. Chicago: University of Chicago Press.
Sag, Ivan. (1992) Taking performance seriously. In Carlos Martin-Vide, ed., *VII Congreso de Languajes Naturales y Lenguajes Formales*.
Sandra, D. and Fayol, M. (2003). Spelling errors as another way into the mental lexicon. In Baayen, R.H. and Schreuder, R. (eds.) Morphological Structure in Language Processing. Berlin: Mouton de Gruyter. 485-514.
de Saussure, Ferdinand. (1916/1959) *Cours de linguistique générale*. English translation by Wade Baskin, New York: Philosophical Library.
Schreuder, Robert, Cristina Burani, and R. Harald Baayen (2003). Parsing and semantic opacity. In E. M. H. Assink & D. Sandra (eds.), *Reading complex words: Cross-linguistic perspectives*, 159–189.
Selkirk, Elizabeth. (1982). *The Syntax of Words*. Cambridge, MA: MIT Press.
Sereno, J., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory and Cognition*, 25, 425–437.
Spencer, Andrew (2013). *Lexical Relatedness*. Oxford: Oxford University Press.
Stump, Gregory. (2001). *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
Talmy, Leonard. (1978). The relation of grammar to cognition: A synopsis. In D. Waltz (ed.), *Theoretical Issues in Natural Language Processing 2*. New York: Association for Computing Machinery. Revised version in Talmy, *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press 2000.
Tomasello, Michael. (2003). *Constructing a Language*. Cambridge, MA: Harvard University Press.
van der Hulst, Harry. (2006). On the parallel organization of linguistic components. *Lingua* 116, 657-688.
Van Valin, Robert, and Randy LaPolla. (1997). *Syntax: Structure, Meaning and Functions*. Cambridge: Cambridge University Press.
Yang, Charles. (2005). *Yearbook of Language Variation*, 5, 333–370.