1. Introduction


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) articulate 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?

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.

---

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 early version. RJ also thanks participants in the Tufts Linguistic Research Seminar (Ari Goldberg, Neil Cohn, Eva Wittenberg, Anita Peti-Stantic, Naomi Caselli, Anastasia Smirnova, Rabia Ergin, Irit Meir, Rob Truswell, and Katya Pertsova), who endured numerous primordial 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.
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 stored in memory.

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 generative systems – 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.

![Figure 1. The Parallel Architecture](image)

How do words fit into Figure 1? All theories agree that a word involves an association between a piece of phonology, a piece of meaning, and a collection of syntactic features such as part of speech, grammatical number and gender, and so on, along the lines of (1). Within the Parallel Architecture, then, a word can be thought of as a small-scale interface rule, linking pieces of semantic, syntactic, and phonological structure.

(1) Semantics: [CAT]
Syntax: N
Phonology: /kæt/

(1) is a stereotypical lexical item: a word with content in all three components, together making up a Saussurian sign. However, the lexicon also contains words 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 chew the fat (‘converse idly’) are another sort of nonstereotypical lexical item: they have internal syntactic structure but noncompositional semantics (and hence within this idiom, chew and fat are not signs).²

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
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. Consider how the word *cats* might be encoded. Semantically, plurality is an operator on CAT; syntactically, it is a feature on a noun that plays a role in agreement; and phonologically, it is an affix.

(2) Semantics: \([\text{PLUR } ([\text{CAT}])]\)
Syntax: \([\text{Npl } N, \text{ pl}]\)
Phonology: /kæt/ /s/

If we extract the contribution of the word *cat* to (2), we arrive at a schema for the English regular plural, more or less in the style of Construction Morphology. This schema contains variables at all three levels, indicated by underlines.

(3) Semantics: \([\text{PLUR } (X)]\)
Syntax: \([\text{Npl } N, \text{ pl}]\)
Phonology: /z/

The variables in (3) can be instantiated by a lexical item such as *cat* (example (1) above) to form the plural noun (2). (We assume that the voicing assimilation of the suffix is a matter of phonology, not morphology.)

A morphologically complex word like *cats* in (2) incorporates a stereotypical mapping between semantics, syntax, and phonology: each part in each component corresponds one-to-one to a unique part in the other components. (We will make this correspondence formally more precise in section 3.2.) Such stereotypical cases conform to the classical notion of the morpheme, in which an identifiable phonological string within a word corresponds to an identifiable part of the word’s meaning. However, the deviations from this one-to-one linking are legion: meaningless case and agreement markers, zero morphology (including conversion), stem allomorphy (including ablaut), portmanteau forms, truncation, infixation, circumfixation, multiple exponence, and reduplication. The ubiquity of such phenomena has led to widespread skepticism toward the morpheme as a viable construct of morphological theory (e.g. Anderson 1992).

Within the PA, such deviations from stereotypical morphemes can be treated as noncanonical links between morphosyntax and phonology. For instance, a meaningless accusative case marker links phonological and morphosyntactic structure, but there is no corresponding semantic structure. This parallels meaningless words in phrasal syntax such as epenthetic *it*. For a different sort of case, a noun like *sheep* (4a) has a plural (4b) whose

---

function (semantics). PA countenances meaning-free lexical items as well as meaningful ones. See Jackendoff 2013 for discussion.

3 Construction Grammar would use the term *construction* for (3); we follow the practice of Construction Morphology (Booij 2010) in calling it a schema, in order to emphasize its difference from procedural rules.
semantics includes plurality and whose morphosyntax includes the plural feature, which is necessary in order to control agreement. In other words, (4b) is just like (2) – except that it has no uniquely identifiable phonological material corresponding to plurality. However, there is no need to introduce a phonological “zero morph” in order to preserve a one-to-one correspondence; this is just a noncanonical correspondence between the levels of structure.4

(4)  

a. Semantics: SHEEP  
Syntax: N  
Phonology: /ʃip/  

b. Semantics: [PLUR (SHEEP)]  
Syntax: [Npl N, pl]  
Phonology: /ʃip/  

These analyses suggest that 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 2. 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.)

Figure 2. Morphology in the Parallel Architecture

To understand these components and their relations, first consider morphosyntax – the internal syntactic structure of words. Inflectional morphosyntax stipulates a language’s dimensions of inflection, creating an n-dimensional array of possible inflected forms: the grammatical features and their range of values, e.g. gender, number, and case for nouns; tense, person, and number for verbs. Derivational morphosyntax stipulates the repertoire of derivational affixes of a language and their effects on syntactic category, for instance, that the -(t)i on suffix of English combines with a verb to form a noun (with semantic consequences that we mention in a moment).

In a syntactic structure, morphosyntax interfaces with phrasal syntax at the level of maximal X0’s, which constitute the largest morphological entities and the smallest syntactic ones. This level includes inflectional features (Anderson 1992), which are visible to both morphology and phrasal syntax. Aside from X0 categories, morphosyntax and phrasal syntax do not share

4 Notice the parallel to idioms: Just as the phonology of irregularly inflected forms cannot be partitioned in one-to-one correspondence with their morphosyntactic features, the semantics of idioms often cannot be partitioned into the meanings of its syntactic constituents. (Which part of ‘converse idly’ comes from chew and which from fat?)
categories: affixes are found in morphosyntax but not phrasal syntax, while phrasal syntax has categories like NP, VP, and S, which with certain exceptions\(^5\) play no role in morphosyntax. (5) illustrates the spheres of responsibility for the components in the VP *finalized the agreements*.\(^6\)

The upper ellipse is the domain of phrasal syntax, including inflection but not derivation; the lower ellipse is the domain of morphosyntax. Notice that phonology is its own level of representation, and so (following the PA) is in neither ellipse.

(5)

![Diagram of spheres of responsibility for the components in the VP](image)

---

Turning to the other components of Figure 2: word phonology concerns the phonological shape of words, including matters such 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 to morphological theory are the interface principles that link phonology to morphosyntactic structure. As pointed out above, morphosyntactic constituents stereotypically map one-to-one to phonological constituents, such that each piece of morphosyntax has a corresponding string of sounds. However, languages are full of variants on the stereotype, such as plural *sheep* in (4b).

Within morphosyntax, inflectional morphology is (for the most part) perfectly regular: every verb and every noun can appear with every combination of appropriate inflectional features.\(^7\) At this level, then, a conventional item-and-arrangement conception of morphology is altogether feasible. 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 words cannot be split cleanly into identifiable morphemes.

On the semantic side of Figure 2, similar considerations obtain. “Word semantics” specifies the possible semantic forms that words can take – the range of “lexical conceptual

---

\(^5\) 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, boîte-sans-soif*, and English *matter-of-factness* and *matter-of-factly*. These can be treated as word-internal phrasal constituents, e.g. [\([N [NP matter of fact] [adj ness]]\), violating the canonical relation of morphosyntax and phrasal syntax. See Meibauer 2007 for some discussion.

\(^6\) (5) is a random example chosen for its combination of inflection and derivational morphology. We notate the morphosyntax of inflection in terms of features on maximal X\(^0\)s, and that of derivation as hierarchical tree structures. There are of course other possibilities, which we do not have space to discuss here.

\(^7\) This corresponds to Corbett’s (2012: 163) Criterion 4 for canonical morphosyntactic features: “Canonical features and their values are distinguished consistently across lexemes within relevant parts of speech”.

---

5
---

---

---
What might be called “morphosemantics” is the interface mapping between morphosyntactic structure and word meanings. Morphological combination can be used to express a heterogeneous collection of semantic functions – matters like 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.

Overall, then, the scope of morphology proper can be seen 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, 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 have called schemas – rather than in terms of procedural rules that apply in serial order to convert an “input” into an “output.” Even the most fundamental phrase structure rules of syntax come to be stated in this fashion. For instance, in place of the traditional phrase structure rule for the English transitive VP (6a), these frameworks substitute a “treelet” (6b) that licenses VP structures.

(6)  
a. VP → V – NP  
b. [VP V – NP]

From the perspective of the previous section, (6b) can be thought of as a schema, listed in the lexicon, that specifies only syntactic structure. Thus, as in Construction Grammar and Construction Morphology, there is complete continuity between words and rules; there are no separate “places” for them. Both are listed in the lexicon in a common format: as pieces of linguistic structure. A lexical item is more wordlike to the extent that its content is completely specific, such as the entry for cat in (1); it is more rule-like to the degree that it contains

8 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.
variables, like the schema for the plural in (3) or, more extreme, the VP schema in (6b), which consists exclusively of variables.

This feature of the PA makes it incompatible with any theory of morphology whose rules build words procedurally, along the lines of “to form the plural of a noun, add /z/ and assimilate voicing” or “to form the past tense of sing, change /l/ to /æ/.” Such approaches include, among others, A-morphous Morphology (Anderson 1992), Paradigm Function Morphology (Stump 2001, Luis this volume), and Distributed Morphology (Halle and Marantz 1993, Siddiqi this volume).

Schemas and procedural rules are sometimes thought to be notational variants, and this is indeed true of the simplest cases. For instance, the VP schema (6b) can be thought of as a “precompiled” output of the VP rule (6a). 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 – rules and lexicon – 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 utterances. In procedural theories, the rules build the utterance 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 are instantiated by other items. There is no inherent order of derivation: structures can be assembled from the bottom up, from the top down, or from left to right. From the point of view of pure linguistic theory, this might make little difference. But from the perspective of the PA, which values connections to psycholinguistics, the difference is considerable. The order of derivation in procedurally based theories bears no relation to the order in which sentences are processed, in either comprehension or production. In contrast, the 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, alluded to in section 1, and which we now dwell on at some length, 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. In the phrasal domain, idioms such as chew the fat are a prominent case: they have canonical syntactic structure but noncompositional semantics. Morphology, of course, presents many types of phenomena that require lexical storage, such as:

- 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.

---

9 This is one important reason why linguistic theory has needed to invoke the competence-performance divide: as a bulwark against psycholinguistic considerations such as how the grammar is to be implemented in processing.
• 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*.\(^\text{10}\) 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 2009 for inflection). And of course texts of any length – even, say, the entire Koran – can be stored in memory without losing internal structure.

Traditional procedural rules are inadequate for describing these phenomena. 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*. Some of these difficulties can be mitigated with some variety of “exception features” (Chomsky and Halle 1968, Lakoff 1971); but some, especially idiosyncratic semantics, cannot (Chomsky 1972).

3.2. **Lexical redundancy rules**

A widely accepted approach to these difficulties is the Lexicalist Hypothesis (Chomsky 1972): regular, predictable patterns belong to syntax, while more idiosyncratic 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 LFG (where it was called Lexical Integrity), HPSG, 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.

Translated into the present notation, a lexical redundancy rule looks formally just like the schemas introduced in previous sections – a piece of linguistic structure containing variables. In

---

\(^{10}\) This is true even taking into account the phonological constraints on the base: it must be monosyllabic and end with an obstruent.
order to illustrate, we begin with (7), the lexical entries for hard and harden. We introduce a coindexation notation, which allows us to specify which parts of each entry correspond to each other and to parts of other entries. (One can think of the coindices as marking the end of association lines.) Subscript 1 connects the three components of hard. But it also connects the same components inside the structure of harden. Subscript 2 connects the three components of the whole word harden. Subscript 3 connects the affix’s phonology with its morphosyntax, where it combines with an adjective to form a verb.\textsuperscript{11}

(7) a. Semantics: HARD\textsubscript{1} b. Semantics: [BECOME (HARD\textsubscript{1})]\textsubscript{2}
Syntax: A\textsubscript{1} Syntax: [v A\textsubscript{1} aff\textsubscript{3}]\textsubscript{2}
Phonology: /hard/\textsubscript{1} Phonology: //hard\textsubscript{1} /əәn\textsubscript{3} /\textsubscript{2}

Extracting the schema for the affix, just as we did for the plural in (3), we arrive at schema (8), which relates the pair in (7) as well as wide/widen, tough/toughen, and so on.

(8) Semantics: [BECOME (MORE X\textsubscript{x})]\textsubscript{y}
Syntax: [v A\textsubscript{x} aff\textsubscript{3}]\textsubscript{y}
Phonology: //…/x /əәn\textsubscript{3} /y

Schema (8) says that an adjective, whatever its meaning (X) and whatever its pronunciation (we omit the phonological constraints mentioned in note 10 above), can be incorporated into a verb that means ‘become more X’ and that has the affix /əәn/ tacked onto the end of the base. The variable co-indices x and y allow this schema to be related to any word such as harden and widen that has a corresponding pattern of co-indexation. On the other hand, the affix /əәn/ itself is constant throughout all instances, so in schema (8) as well as in harden (7b) it is encoded with a constant numerical co-index 3.

Schema (8) is intended to fulfill the function of a so-called lexical redundancy rule. It expresses a relation among items stored in the lexicon – it does not “generate” them, and it is not expected to apply across the board. If an adjective happens to have a corresponding verb that falls under (8), this is an advantage, in a sense we will discuss in section 4. But if there is no such pairing – for instance loud does not have a counterpart *louden – that is also satisfactory.

But before asking what benefit a nonproductive schema like (8) can confer, we need to consider its status in relation to regular, productive schemas like the plural (3).

\textsuperscript{11}The notation for the phonology of harden perhaps calls for further explanation. We use paired slashes //…// to set off phonological strings that can be linked to constituents in morphosyntax and semantics. Thus the string /hard/ is linked to the adjective and the meaning HARD, and the string /əәn/ is linked to aff. However, in addition, the entire string is linked to the whole morphosyntactic and semantic constituent; this is notated by the outermost pair of slashes and the coindex 2.
3.3. Distinguishing productive from nonproductive schemas

The problem is that fully productive schemas, such as the plural, and nonproductive schemas, such as *A-en, are formally exactly parallel: they are pieces of phonological, syntactic, and semantic structure containing a variable on each level. How does the grammar say that one is productive and the other is not? The Lexicalist Hypothesis would say that the plural is “in the grammar” while *A-en is “in the lexicon,” and hence that they are actually of quite different status. In the Parallel Architecture, though, the “grammar” itself is “in the lexicon,” so this distinction is not available, and the problem remains.

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 rather on its variable. For instance, the variables in the plural schema (3) would be marked productive and those in the *A-en schema (9) would be marked nonproductive. If it proves desirable, such marking could encode a graded degree of productivity, capturing the intuition that a schema’s degree of productivity amounts to the degree of openness of its variable to new lexical material – the willingness of speakers to extend the schema to new instances (as, among others, in Aronoff 1976; Lieber and Baayen 1991; Baayen 1993; see also Hüning this volume and Bauer 2001 for discussion).

Again, this looks superficially like a notational variant. But it leaves an interesting loophole: 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 (9a-d). The italicized words name the type of geographical feature.

(9) 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 want to name a mountain for Morris Halle, we have no hesitation in calling it Morris Mountain or Mount Morris; if we want to name a mountain range for him, it will be the Morris Mountains (like (9c)). 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 and *the Mountain Morris are impossible. Hence the schemas for these patterns have one variable of each type, as in (10). 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 (10c,d) are constants.

(10) a. N N [e.g. Loon Lake]
b. N N [e.g. Mount Washington]
c. the N N [e.g. the Lehigh River]
d. the N of 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. Nor is it very natural within the Lexicalist Hypothesis, in which the name is inserted “in the grammar” but the choice of pattern, determined by the type of geographical feature, is “in the lexicon.”

The patterns in (9)-(10) offer another kind of evidence for the PA’s treatment of morphology and against a strict division between lexicon and grammar. On one hand, (9a,b) look like compounds. For example, they can be preceded by adjectival modifiers: beautiful Arrowhead Lake, forbidding Mount Everest. On the other hand, (9c,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. (9d) 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 (10) are in the lexicon: (10a,b) are morphosyntactic; (10c), in which the two nouns still form a compound, is a mixed morphosyntactic and phrasal schema; and (10d) is purely phrasal.

Here is where things stand at the moment, then: Nonproductive schemas, with closed variables, capture patterns among stored lexical items. Their instances must be listed in the lexicon, and their coinages are noticed as novel. Furthermore, they tolerate instances with idiosyncratic properties such as the noncompositional semantics of recital, the irregular phonological relation of destroy and destruction, and the absence of a root as in impetuous. In contrast, productive schemas, with open variables, allow free online combination with anything that satisfies the conditions of the variable, as in for instance novel plurals such as wugs and freely chosen transitive verb phrases such as slice some pastrami.

However, the distinction between open and closed variables is not quite so stark. Productive schemas actually have two distinct functions. Linguistic theory has for the most part concentrated on their generative function, their use in creating new expressions. But they can also be used to capture patterns among stored lexical items, just like nonproductive schemas. We might call this use their relational function. 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 fall under the compound schema in its relational role. Similarly, the productive schema for plurals applies in its relational function 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, and so it has to be accounted for by the plural schema in its relational function. More generally, experimental evidence suggests that highly frequent regular plurals are likely to be stored rather than generated online (Baayen 2007; Nooteboom, Weerman, and Wijnen 2002), so that cats is more likely to be stored than, say, coelacanths. In each of these cases, then, the stored items fall under the very same schemas as

---

12 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.
the ones that are actively composed. The schema is just used relationally rather than generatively.

This dual use of productive schemas is not limited to morphology. In phrasal syntax, the schema for transitive VPs (6b) applies not only generatively, to construct novel VPs online, but also relationally, to license stored idioms with non-compositional semantics such as *chew the fat* and *kick the bucket*. Moreover, nonproductive schemas – schemas that are used only relationally – turn up in phrasal syntax, although far less prominently than in morphology. The geographical place names in (9c,d) are one case; for another, the NPN construction (*day after day*) occurs with only a limited class of prepositions (11a); and the unusual determiner formula in *what a job*, which occurs only with a restricted set of words preceding the indefinite article (11b). Culicover 1999 calls such patterns “syntactic nuts”; their instances have to be learned one by one.

(11)  a. NPN: day after day, week by week, face to face; *person from person, *gun beside gun

b. [Det X a]: what a job, such a job; *who a professor, *where a city

This observation leads to an interesting change of perspective: one can think of productive schemas as just like nonproductive ones, except that in addition, they allow free online composition with their variables. In other words, productive schemas are “ordinary” schemas that have “gone viral.” This observation is possible only in a constructional framework like Construction Grammar/Morphology or the PA. And it turns on its head the notion that the study of language should focus primarily on productive phenomena. As suggested in section 1, the focus should be equally if not more on the relationships among items in the lexicon, that is, the texture of linguistic knowledge.

3.4. Are nonproductive schemas necessary?

We take it as given that productive schemas or their equivalent are necessary in order to account for the construction of an unlimited number of sentences. However, an important question for morphological theory is whether there actually are such things as nonproductive schemas. We have just pointed out some interesting consequences of assuming that they are necessary. However, this assumption is not universal. 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. Similarly, Spencer (2013, 3), speaking of nonproductive patterns, says: “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).”

---

13 This fact was noticed in Jackendoff 1975, but it made no sense at the time, as the theory of the time made a strict distinction between syntactic rules and lexical redundancy rules.

14 Could there be schemas that have only the generative function and lack the relational function? No, because one can always store newly generated instances of a schema, which then fall under the schema’s relational role.
There are certainly 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 (12) appear to be unique in their phonological relations to each other; there is no more general rule or schema that relates them.

(12)  
   a. surgeon/surgery  
   b. bomb/bombard  
   c. hate/hatred  
   d. laugh/laughter  
   e. humble/humility  
   f. bequeath/bequest  
   g. Glasgow/Glaswegian

A more questionable case is horror/horrify/horrid/horrible/horrific/horrifying. 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 how analytically they treat the language, and in how eager they are to form schemas (unconsciously, of course).\footnote{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.}

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 number 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 and other constructionist theories, a hypothesis has 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). 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 is useful. (For more discussion, see the next section and Jackendoff and Audring in preparation.)

In the PA 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 No Rule and 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 unproductive, this does not mean it is wrong – the observed pattern of sameness and difference remains valid. If, on the other hand, a schema is found to be productive, it does not have to relinquish its status as a lexical redundancy rule. Rather, as suggested in section 3.3, it still retains the function of capturing generalizations among stored items.

Nonproductive schemas can also be 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.

4. Formalizing lexical relations

4.1. Inheritance with impoverished entries and full entries

In section 3.2, we spoke of the \([\nu \ A – en]\) schema “applying” to harden. We now have to ask what this locution can mean within the PA approach. How are hard and harden related, and what does the schema have to do with it?

We have already rejected one standard position: that harden is derived from hard by a procedural rule along the lines of “to form a verb that means ‘become more A’, add /əә/ to A” – whether such a rule is “in the grammar” or “in the lexicon.” Such a derivation may in some cases account for how speakers improvise new forms that they haven’t heard, or how the verb entered the language at some point in the historical past. However, harden is far from an improvisation, and speakers are largely unaware of the history of words. Within their minds, both hard and harden are stored. Hence the lexicon must have a way of relating them, and the schema must somehow be involved in stating that relation.

A common position (e.g. Booij 2010) is that hard, harden, and \([\nu \ A – en]\) are related through an inheritance hierarchy, as illustrated in (13). 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 and from the schema, while whiten inherits from white and from the schema.

---

16 -- for instance, when JA asked RJ if he was going to dutchify his bike, i.e. equip it with handlebars in the Dutch style.
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:

\[
\text{A} \text{hard } [V \text{ A-en}] \text{ Awhite} \\
[V \text{Ahard en}] [V \text{Awhite en}]
\]

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 in preparation). 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. Experimental evidence does not bear this prediction out: if a complex item is more frequent than its base – for instance eyes is more frequent than eye – it is accessed as quickly as its base (e.g. Baayen, Dijkstra & Schreuder 1997).

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

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.

\[
\text{A} \text{hard} [V \text{ A-en}] \\
[\ldots]
\]

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 (15).

\[
[A \text{N – ous}] \\
[\text{impet} . ]
\]
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
suggests 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.”
This assumption is what Langacker 1987 calls the “rule-list fallacy”: that if something can be
described by rules, it ought not to be listed as well. 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
themetic roles redundantly, through word order, case marking, and verb agreement. Further
afiel, the visual system has at least five partially redundant mechanisms for depth perception:
lens accommodation, eye convergence, stereopsis, occlusion, and perspective (Marr 1982).

Given these precedents, we find it more attractive to adopt a full-entry theory, in which
lexical items are encoded in their entirety, even where redundant. For instance, the full-entry
theory says that the lexical entry of harden is like (7b), with its full structure, rather than like
(14), 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
and indeed by Saussure (1915, 133).

While the notion of motivation is intuitively attractive, not enough is known at the
moment about neural storage and computation to cash it out in brain terms. Nevertheless, we
will adopt it heuristically, notating it in terms of shared and coindexed structure.

To illustrate, we return to our treatment of harden, repeated here.

(16) a. Semantics: HARD₁ b. Semantics: [BECOME (HARD₁ ) ]₂
Syntax: A₁ Syntax: [ v A₁ aff₃ ]₂
Phonology: /hard/₁ Phonology: / /hard/₁ /əә n/₃ /₂

(17) Semantics: [BECOME (MORE Xₙ)]ₚ
Syntax: [ v Aₙ aff₃ ]ₚ
Phonology: / /ₚ /əә n/₃ /ₚ

¹⁷ 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.
The shared structure is encoded in the coindices. Consider for instance the subscript 1. Within (16a), it ties together the semantics, syntax, and phonology of the word *hard*. It also appears in (16b), tying together constituents of *harden* in semantics, syntax, and phonology. At the same time, though, it links these constituents to the corresponding parts of (16a), thereby identifying the structure shared between the two entries. Similarly, subscript 3 picks out structure shared between *harden* and schema (17): the affix and its phonological realization. Finally, as suggested above, the variable subscripts *x* and *y* in (17) indicate structure shared with anything that has parallel structure – i.e. all the instances of the schema.

An immediate advantage of the notation in (16)-(17) is that it enables us to pinpoint the regions of similarity among items. The standard notation for inheritance, as in (13), leaves it somewhat vague as to what specific parts of items are inherited from what (though, of course, there are more sophisticated models, e.g. Brown (this volume)).

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, the general filling in the particular. For instance, a complex item such as *harden* inherits from its base (*hard*) and the schema [v A – en]. However, we observed above that in acquisition, schemas have to be constructed from the bottom up, on the basis of some number of previously stored examples. Hence in a sense a schema is motivated by its instances, rather than (or in addition to) the other way around. This conundrum does not arise with the coindexing notation in (16)-(17), which is not inherently directional. We do not have to decide whether (17) is motivated by (16b) or the other way about: 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.

(18) shows how such cases can be treated in the coindexing notation. The semantics of (18b), coindexed 5, is shared with part of (18a), and the phonology of (18a), coindexed 4, is shared with part of (18b). (Coindex 6, 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.

---

18 It should be mentioned that this problem is not specific to the full-entry theory. It is if anything more severe for derivational and impoverished-entry theories.
(18)  a. Semantics: [PERSON WHO [MURDERS POLITICIAN]_5 ]_4  
Syntax: N_4  
Phonology: /əәsæsən/ 

b. Semantics: [MURDER POLITICIAN]_5  
Syntax: [V N_4 – aff_6 ]_5  
Phonology: //əәsæsən/ /eyt/ 

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. One way to express this is as (19a), where the morphosyntax of the base is blank: this part of the word is in effect unparsed in morphosyntax, although it does have phonology. There are hundreds of such –ous words (e.g. scrumptious, curious, impetuous, supercilious, meticulous), in addition to those with a noun base such as joyous (19b) (where coindex 10 links to the noun joy). So there ought to be a schema expressing this possibility, along the lines of (19c). The angle brackets in (19c) allow for a base that is either a noun or unparsed.

(19)  a. Semantics: BEAUTIFUL_7  
Syntax: [A – aff_8 ]_7  
Phonology: //gordʒ/ /əs/ 

b. Semantics: [Property <PERTAINING TO (JOY_10)>]_9  
Syntax: [A N_10 – aff_8 ]_9  
Phonology: //dʒoy/ /əs/ 

c. Semantics: [Property <PERTAINING TO (X_2)>]_w  
Syntax: [A <N> aff_8 ]_w  
Phonology: // /əs/ 

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

A more extreme case is found in the words in (20).
(20) 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. (21a) shows an entry for trillion and (21b) is a possible –illion schema.

(21) a. Semantics: $10^{12}_{11}$  
   Syntax: Numeral$_{11}$  
   Phonology: /trillion$_{12}$/$_{11}$

   b. Semantics: LARGE NUMBER
   Syntax: Numeral
   Phonology: /___/illion$_{12}$/

(21) 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.

   Going still further in this direction, we find phonaesthemes such as flitter, glitter, jitter, and pitter-patter; gleam, glow, glisten, glint, and (again) glitter. In these cases only a somewhat vague semantic feature ties the instances together. There is certainly no morphosyntax. Yet there might well be a schema that captures the similarity among them, which like (21b) correlates only phonology and semantics.

   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 the PA’s overall outlook on the character of the lexicon.

4.4. Sister relations

The non-word bases in 4.3 bring us to another case of motivation that poses a serious problem for inheritance. This problem manifests itself in 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. Rather, we would like to be able
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.\(^\text{19}\)

The coindexing notation makes it straightforward to express such sister relations, for instance as in (22). The sisters share part of their semantics (script 13) and part of their phonology (script 17). But they do not share their affixes: subscript 15 is coindexed with the –tion schema (not shown) and subscript 16 with the –tious schema. Finally, since *ambi is not a word and has no part of speech, we leave its morphosyntax blank, as we did for gorgeous in (19a). Crucially, unlike hard/harden (16a,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.

(22) a. Semantics: DESIRE\(_{13}\)  
    Syntax: \([N - \text{aff}15]_{13}\)  
    Phonology: //æmbi/\_17//ʃəә/\_15/\_13\  
  b. Semantics: [HAVING (DESIRE\(_{13}\))]\(_{14}\)  
    Syntax: \([A - \text{aff}16]_{14}\)  
    Phonology: //æmbi/\_17//ʃəәs/\_16/\_14\  

4.5. Sister relations among schemas

As mentioned in section 3.4, some sister relations, such as surgeon/surgery, 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 (23):

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

To express this generalization, we have to establish a relation not between two sister words as in (22), but between two sister schemas: the –ism schema for ideologies and the –ist schema for their adherents.\(^\text{20}\)

(24) a. Semantics: IDEOLOGY\(_x\)  
    Syntax: \([N - \text{aff}18]_{x}\)  
    Phonology: //\ldots//y//ɪzm/\_18/\_/\_\  
  b. Semantics: [ADHERENT (IDEOLOGY\(_x\))]\(_z\)  
    Syntax: \([N - \text{aff}19]_{z}\)  
    Phonology: //\ldots//y//ɪst/\_19/\_/\_\  

Let us unpack this. (24a) says that there can be nouns that denote an ideology and that end with the affix –ism (coindex 18). The variable coindex \(x\) ties the three components together. (24b) says that there can be nouns that denote the adherents of an ideology and that end with the affix –ist (coindex 19); the variable coindex \(z\) ties these three components together.

\(^{19}\) This sort of case played a central role in Jackendoff’s (1975) arguments about the nature of lexical redundancy rules.

\(^{20}\) In turn, the –ist schema for adherents of an ideology is a special case of the –ist for occupations such as linguist, botanist, dentist, and trombonist, each of which has a different phonological relation to the corresponding activity (linguistics, botany, dentistry, and [playing] trombone).
So far these schemas are just like the –ous schema in (17). But what makes them more interesting is that IDEOLOGY in (24a) is linked to IDEOLOGY in (24b) by the variable coindex \(x\), and the part of the phonology in (24a) that precedes the affix is linked to the corresponding part in (24b) 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 (2010: 31-36, Booij & Masini 2015) calls this configuration a second-order schema (see also Nesset 2008 and Kapatsinski 2013). It can be considered a constrained form of analogy, e.g. pacifism and altruism are to (24a) as pacifist and altruist are to (24b). As Booij points out, second-order schemas offer a way to describe paradigmatic inflectional patterns: each inflectional form in the pattern is represented as a schema, and the entire collection of schemas is linked together by shared variable coindices. They are also useful in describing ablaut and other forms of stem allomorphy, as well as truncations such as Michael/Mike. Space precludes providing more details here; the reader is referred to Jackendoff and Audring (in preparation).

Booij notates a second-order schema as a special relation between schemas: 
\(< \text{Schema A} > \approx < \text{Schema B} >\). In the present approach, second-order schemas are a generalization of sister relations. Where the sister relation in (22) links numerical coindices across two words, the second-order schema in (24) 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 a theory of morphology grounded in the outlook of the Parallel Architecture. The fundamental goal is to describe what a speaker stores and in what form, and to describe how this knowledge is put to use in constructing novel utterances. A basic tenet of PA is that knowledge of language is segregated into phonological, syntactic, and semantic/conceptual structures, plus interfaces between them that enable sound, morphas syntactic 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 or stereotypical 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 Grammar, the Parallel Architecture 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. A morphological theory consistent with this approach must likewise state morphological patterns in terms of declarative schemas rather than procedural or “realizational” rules.

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. On the other hand, a substantial proportion of morphological patterns are not productive. They can be described in terms of schemas that are formally parallel to those for productive patterns, except that they have only the relational function. We have argued 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, words and schemas, and 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 (in preparation), for instance zero morphology, stem allomorphy, blends, truncations, and inflectional classes. However, even in the present chapter, we have touched 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


Siddiqi, Daniel. (this volume). Distributed Morphology.