Shifting paradigms: gradient structure in morphology

Jennifer B. Hay[1] and R. Harald Baayen[2]

[1]Department of Linguistics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand
[2]Interfaculty Research Unit for Language and Speech, Radboud University Nijmegen and Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands

Morphology is the study of the internal structure of words. A vigorous ongoing debate surrounds the question of how such internal structure is best accounted for: by means of lexical entries and deterministic symbolic rules, or by means of probabilistic subsymbolic networks implicitly encoding structural similarities in connection weights. In this review, we separate the question of subsymbolic versus symbolic implementation from the question of deterministic versus probabilistic structure. We outline a growing body of evidence, mostly external to the above debate, indicating that morphological structure is indeed intrinsically graded. By allowing probability into the grammar, progress can be made towards solving some long-standing puzzles in morphological theory.

Introduction

According to the symbolic view [1–5], morphological structure is combinatorial. A past-tense form like *walked* consists of two discrete elements, the MORPHEMES (see Glossary) *walk* and *ed*, which have independent entries in the mental lexicon. A deterministic rule combines these morphemes into the complex structure *walk+ed* and compositionally derives its meaning from the meanings of the parts. The resulting complex forms are not (in more recent versions of the theory, need not be) stored in memory. By contrast, irregular past tense forms like *came* are stored in memory, and are not rule-governed.

According to the subsymbolic view, both regular and irregular forms can be accounted for with subsymbolic networks. Connectionist networks can be trained to map present tense forms (*come/walk*) onto past tense forms (*came/walked*) without requiring overt symbolic rules and without making an a priori distinction between regular and irregular verbs [6–8]. In this approach, morphological structure is inherently probabilistic, experience always leaves traces in memory irrespective of irregularity, and the meanings of complex words can be affected in subtle ways by similarity [9,10].

According to the symbolic model with entries and rules, discrete non-probabilistic combinatorial structure lies at the heart of morphology. According to the subsymbolic approach, morphological structure is fundamentally non-discrete. Instead, morphological structure emerges from the statistical regularities that characterize the forms and meanings of words. In this view, morphological structure is inherently graded. As demonstrated by recent papers in this journal [4,5,11], the two sides seem to be locked in stalemate.

Whether morphological structure is graded is therefore a contentious topic. However, there is convergent evidence supporting the gradience of morphological structure. A substantial part of this evidence comes from studies that are external to the above debate, and that are not committed to the connectionist subsymbolic perspective. This review outlines this body of work, and demonstrates that a probabilistic approach to morphological structure can bring insight to some long-standing issues in morphological theory.

In this review we separate the questions of subsymbolic versus symbolic implementation on the one hand, from discrete versus gradient structure on the other. The results summarized could potentially be modelled both

---

Glossary

Formative: a subpart of a complex word serving a syntactic and/or morphological function (e.g. *jumping, helmsman*)

Inflectional morphology: deals with the changes in a word's form that mark its function in the sentence (e.g. *walk+ed, jump+s*). Can be distinguished from **derivational morphology** in that derivation creates new words (e.g. *hand+ful, climb+er*).

Inflectional paradigm: the set of inflected forms for a particular word (e.g. *ring, rang, rung, rings, ringing*). Table 1 gives an example from Estonian. Inflectional paradigms are one of a range of important paradigmatic relationships, as illustrated in Figure 1.

Juncture: the location of the boundary between two linguistic elements - here, the juncture between meaningful subparts of a complex word. Complex words can vary considerably in how saliently juncture is marked.

Junctural phonotactics: the sequence of sounds which spans the boundary between subparts of a complex word. The more likely this sequence is to occur inside simple words (e.g. *insincere*, c.f. *tinsel*), the less salient the boundary becomes.

Morpheme: a theoretical construct referring to the smallest meaningful unit of language. A morpheme can be a simple word (e.g. *cat*), or part of a complex word (e.g *goodness* consists of two morphemes, *-good* and *-ness*).

Opaque: semantically opaque words are words whose meaning is not predictable from the meanings of their constituents (e.g. *hand+some*). (cf. Transparent)

Productivity: the degree to which an affix can be used to create new words.

Relative frequency: the ratio of the frequency of a word (e.g. *swiftly*) and that of its base form (*swift*).

Transparent: semantically transparent words are words whose meaning is predictable from the meanings of their constituents (e.g. *hand+ful*).

Word and Paradigm Morphology (WPM): an approach to morphology in which the morpheme is dispensed with. The focus is on full words, and on the nature of the relationships between these words.

Corresponding author: Hay, J.B. (jen.hay@canterbury.ac.nz).
Available online 28 April 2005

by symbolic and non-symbolic approaches. However, they resist modelling by strictly deterministic, non-probabilistic approaches.

Graded structure in morphology

Traditionally, morphological theory posits a categorical distinction between simple words (e.g. *govern*) and complex words (e.g. *government*). However, people's behavior in experimental tasks is anything but categorical. Individuals can rate affixed forms consistently on a scale from unaffixed to affixed, and can assess which member of a pair of complex words is more complex (e.g. *settlement* is reported as 'more affixed' than *government*) [12–14]. This suggests that morphological complexity is not a binary category. In addition, similarity judgements between affixed forms and their bases are continuous, with no clear division between semantically compositional, TRANSPARENT forms (e.g. *leader*) and semantically non-compositional, OPAQUE forms (e.g. *dresser*) [9,15]. Different degrees of semantic transparency are reflected in degrees of priming [9,15,16], and graded priming effects are also observed with different degrees of phonological or orthographic similarity [17,18].

Supporters of discrete models of morphological structure may argue that gradient behaviour in experimental tasks reflects gradience in processing, or in response strategies, but not in underlying structure. This interpretation becomes problematic in the light of evidence (reviewed below) that gradience is also reflected in speech production, and constrains morphological processes such as affix-ordering. The clearest interpretation of the combined evidence from speech perception and speech production is that morphological structure is inherently graded.

But how can structure be graded? If we decompose *walked* into the morphemes *walk* and *ed*, haven't we assigned a discrete, deterministic decompositional structure to *walked*? It is certainly difficult to see how morphological structure might be graded as long as the morpheme is viewed as the cornerstone of a morphological system which consists of morphemes and rules operating on these morphemes.

However, there are other morphological theories which do allow a graded view of morphological structure. These theories take the position that the morpheme (defined as the minimal structural unit combining form and meaning) is a highly problematic theoretical construct [19–21]. For instance, in many languages of the world, one finds FORMATIVES that in no way participate in a semantic combinatorial system. An example from the Estonian case system is shown in Table 1.

In theories such as WORD AND PARADIGM MORPHOLOGY (WPM) [23] the morpheme is dispensed with. Full words are viewed as the basic units in the lexicon. The degree to which *ed* is 'present' in *walked* depends on the amount of analogical support from other words in the lexicon occupying similar positions in the inflectional paradigm (e.g. *thanked*, *warmed*). Because structure 'exists' in WPM only to the extent that it is supported by exemplar-driven similarity across paradigms, WPM offers a perspective on

Table 1. The inflectional paradigm of singular and plural case endings of the Estonian noun *jalg* ('foot') (simplified after [22])

Case[a]	Singular	Plural
Nominative	'jalg	'jalad
Partititive	'jalga	'jalgasid
Genitive	'jala	'jalgade
Illative	'jalasse	'jalgadesse
Inessive	'jalas	'jalgades
Elative	'jalast	'jalgadest
Allative	'jalale	'jalgadele
Adessive	'jalal	'jalgadel
Ablative	'jalalt	'jalgadelt
Translative	'jalaks	'jalgadeks
Terminative	'jalani	'jalgadeni
Essive	'jalana	'jalgadena
Abessive	'jalata	'jalgadeta
Comitative	'jalaga	'jalgadega

[a]Case forms in Estonian have functions similar to those of prepositions in English. Most of the singular case forms are built on the genitive singular, most of the plural case forms are built on the partitive singular. But there is no corresponding dependence in meaning. Data such as these have led many morphologists to abandon the structuralist notion of the morpheme as the basic morphological unit that would combine form and meaning in an incremental combinatorial system.

linguistic cognition in which morphological structure is inherently graded.

If the word rather than the morpheme is the basic lexical unit, one would expect that all words which have been encountered are stored in long-term lexical memory, irrespective of whether they are simple or complex, and irrespective of whether they are regular or irregular. The hypothesis of storage of full forms ([23] see also [21,24]) recently has received extensive experimental support [25–29]. These experimental results are in harmony with the more general view in memory research that any experience leaves a memory trace, and that, as phrased by Landauer ([30], p. 493), we should not be looking for models and mechanisms that produce storage economies, but rather models 'in which marvels are produced by profligate use of capacity'.

WPM's claim that only full words have representations in the lexicon is arguably too strong. Stems and affixes may well develop their own representations. Even so, such representations probably depend for their continuing existence on the graded support they receive from paradigmatic analogy.

This is an attractive alternative to the entries plus rules model only to the extent that the notion of paradigmatic analogy receives empirical support.

Experimental evidence for paradigmatic analogy

In addition to the extensive evidence for the storage of full-forms, there is a growing body of evidence in the literature supporting the hypothesis that several paradigmatic relations (see Figure 1) characterize lexical representation and co-determine lexical processing.

Recent experimental work shows that the probability distribution of all the distinct forms in a word's INFLECTIONAL PARADIGM affect lexical processing [31–33]. Derived words and compounds also entertain paradigmatic relations, they form morphological families through shared stems (*worm*, *wormy*, *ringworm*, *woodworm*). The size of a word's morphological family has emerged in recent years as an independent predictor of lexical
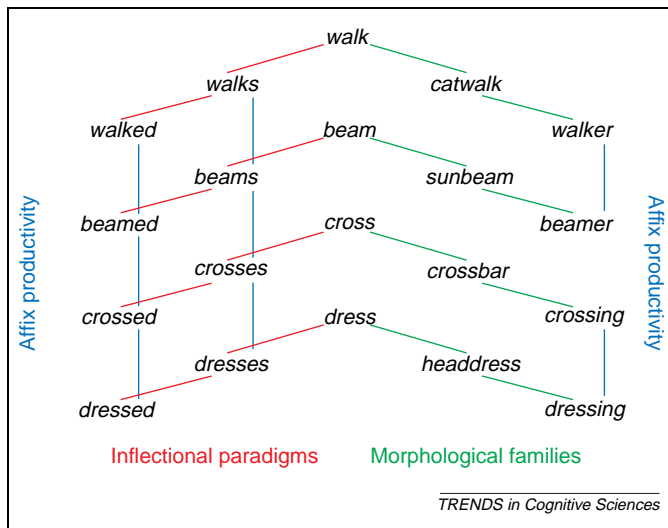
Figure 1. Examples of paradigmatic lexical relations in English. Relations between inflected variants (inflectional paradigms) are shown in red, relations between morphologically related compounds and derived words (morphological families) are shown in green, and relations between words sharing the same affix are shown in blue. Affixes that occur across many words are described as productive.

decision and word naming latencies as well as of subjective frequency ratings in several typologically unrelated languages [34]. These paradigmatic effects show that the relations between words need to be taken into account, and are therefore at odds with the entries plus rules model.

Paradigmatic analogy is also crucial for understanding so-called rule-less morphology. In Dutch, the traditional analysis of the regular past tense required exception marking for a large subset of verbs. However, exemplar-driven paradigmatic analogy can not only explain the distribution of regular past tense formatives (obviating the need for exception markers), but also explains why and where native speakers do not follow the norms of the standard language, that is, where paradigmatic pressure is leading to language change. Furthermore, the strength of the analogical pressure is also reflected in production of the regular past tense [35,36].

Paradigmatic analogy also co-determines the interpretation of novel compounds [37], and guides the use of the

Figure 2. The whole takes precedence over the parts [48]. The monkey wrench representing the horizontal bar of the A has blurred edges ('well-formed junctural phonotactics'), which make it less salient as a part in the whole compared with the diagonal ring spanners with fully distinct edges ('low-probability phonotactic junctures').

otherwise mysterious interfixes (formatives such as the **s** in *helmsman*) which are rare in English compounds but quite productive in German and Dutch compounds [38,39].

A practical example of the importance of paradigms comes from the design and teaching of spelling systems. One of the enigmas of the orthographies of French and Dutch is why even highly educated writers invariably commit morphological spelling errors, even though the spelling rules for complex words are supposed to be simple and transparent. The problem is that analogical memory interferes, allowing inappropriate but more frequent homophones in the paradigm to take precedence [40].

The whole and the parts

We have reviewed evidence for the storage of full words, and the existence of paradigmatic relations between them. However, the parts of complex wholes can also be active during production and comprehension.

There is an extensive literature demonstrating that paradigmatically supported partial matches co-determine lexical processing to different degrees, depending on their distributional characteristics. Apart from partial matches that are themselves words, such as *walk* in *walked* [41,42], there are various other kinds of paradigmatically supported partial matches: bound stems such as *ject* in *inject* [12,43], affixes such as *ed* in *walked* [42,44,45], and 'phonaesthemes' (e.g. *fl* in words pertaining to liquid such as *flow*, *float*, *flood*) [46]. Priming studies show that the combination of overlap in form and overlap in meaning leads to facilitation that tends to exceed the facilitation obtained for form or meaning alone [47]. Interestingly, such superadditive facilitation is obtained not only for 'bona fide' complex words like *walked*, but also for *flow* and *float* [46], and *float* and *boat* (M.J. Pastizzo, unpublished dissertation, State University of New York at Albany, 2004), for which no decomposition into a sequence of discrete morphemes is possible. To account for these graded effects of morphological structure, the morpheme is too coarse a theoretical notion.

It is an open question to what extent bound stems, affixes, and phonaesthemes develop independent form and/or meaning representations. Experimental evidence is often interpreted as supporting 'decomposed' morphological representations (see for example [41]). However, the inference that effects observed for shared lexical structure reflect independent representations, although possible and attractive in its simplicity, is logically not compelling. We are inclined to think that, although independent representations might indeed develop, they depend for their existence on the degree of continuing probabilistic support received from paradigmatic analogy.

Given the combined evidence for full forms and their parts, a question arises about how the whole relates to its parts. For comprehension, it is clear that, as elsewhere in cognition [48], the whole often takes precedence over its parts. Figure 2 illustrates this point by means of an analogy with letter perception. The percept in Figure 2 is that of the letter **A**, albeit an **A** composed of tools. The 'meanings' of the tools themselves do not contribute to the 'meaning' of the letter they form. Many complex words similarly contain elements that themselves do not

contribute to the meaning of the whole, although they have their own meanings (e.g. *corn* in *corner*, *dress* in *dresser*, *bone* in *trombone*, and the partitive singular in the plural case endings in Estonian illustrated in Table 1). Such spurious elements are co-activated during the comprehension process, but the meaning that normally becomes available for further processing is that of the whole [9,49–51].

Many studies have attempted to clarify factors influencing the relative contribution of the whole versus the parts [52–54]. One factor affecting this balance is the ratio of the frequency of the word and that of its base. The more frequent the complex form is relative to its base (e.g. *illegible* is more frequent than *legible*), then the more salient the whole is, relative to its parts. Affixed forms which have a high frequency relative to their bases are rated less complex than forms which are less frequent than their bases, and they are significantly more prone to semantic drift [13,14]. Furthermore, such forms tend to contain fewer phonetic cues to JUNCTURE.

A word's JUNCTURAL PHONOTACTICS concern the probability of the sequence of sounds spanning the juncture between its parts. Low probability, ill-formed, junctural sequences create sharper boundaries and more salient parts (e.g. **in**h*umane* – [**nh**] never occurs in simple words in English). Words with higher probability phonotactics across the morphological boundary (e.g. **in**s*incere*; c.f. *ti***ns**el) have less salient parts [14,44]. With reference to Figure 2, we could say that the monkey wrench representing the horizontal bar of the A has blurred (well-formed) junctures, which makes it less salient as a part in the whole than the diagonal ring spanners, which meet in a more distinct (less well-formed) juncture.

The more the parts 'stand out' in the whole, the stronger the paradigmatic relations that the whole entertains. Affixes represented by more words which are infrequent relative to their bases, and which contain low probability phonotactics, are not only the most likely to be more highly segmentable and to develop stronger independent representations, they are also more readily available for use in new words; that is, they tend to be more PRODUCTIVE.

Morphological productivity

Some affixes (e.g. *-ness* as in *sadness*) are more likely to be used to create new words than others (e.g. *-th* as in *warmth*). The suffix *-ness* is said to be PRODUCTIVE, and *-th* to be UNPRODUCTIVE. Although there is some discussion as to whether an affix is ever truly and totally unproductive [55], most morphologists agree that affixes do display very different degrees of productivity [56].

An initial challenge to understanding the source of these degrees of productivity rests with finding a measure of productivity itself. Measures which formalise the notion of degree of productivity in terms of conditional probabilities that go back to Turing are now available [55,57]. These measures provide tools which can rank affixes according to different aspects of productivity, and have opened the door to studies probing the question of the source of differences in productivity. Whereas affixal productivity cannot be straightforwardly predicted by an

affix's frequency of use, it can be predicted from the degree of paradigmatic support that the affix receives. The reason that simple frequency counts fail is that not all words 'contain' the affix to the same degree. The relative salience of the whole and the parts, as gauged by their relative frequencies and junctural phonotactics, are significantly correlated with affixal productivity. In other words, the degree of productivity of an individual affix is co-determined by the degrees to which the various words containing that affix (its affixal paradigm; see also Figure 1), are morphologically complex [14,58].

Affix-ordering

The hypothesis that morphological structure emerges gradiently from paradigms allows considerable insight into restrictions on English affix-ordering. Most languages have restrictions on the order in which affixes can occur with respect to one another. In English, for example, the affix *-ity* is never attached to the affix *-less*. The nature of the restrictions on affix-ordering has been a long-standing debate. One common approach to the problem has been to characterize affixation as occurring on different 'levels', with level 1 affixes attaching before level 2 affixes. The ungrammaticality of *cluelessity* then follows from a level 2 affix (*-less*) inappropriately nested within a level 1 affix (*-ity*). Many languages have received level-ordered treatments of affix-ordering (see [59] for a review).

The level-ordering perspective has received much criticism [60–62]. One reason for this is that there tend to be restrictions on ordering among affixes within a single level, which a level-ordered account cannot capture [60,61]. Although many contemporary theories maintain a distinction between level 1 and 2 for independently motivated phonological reasons, most have abandoned an explicit level-ordered account of affix-ordering [63].

Interestingly, level-ordering achieved some success, not because there are actually two discrete levels, but rather because there is a systematic relationship between degree of structure and ordering restrictions. The generalization turns out to be that as we move outwards from the stem in a multiply suffixed word, the suffixes encountered are progressively less 'fused' with their host in terms of RELATIVE FREQUENCY and junctural phonotactics [14,64]. In other words, morphological structure fades as we move in towards the stem, a phenomenon that has been observed for inflection as well [65]. As shown in [66], there is a remarkable isomorphism between this graded structural constraint and the notional constraints governing affix selection.

Phonological and phonetic implementation

Languages display strong preferences for formatives in a paradigm to be highly similar - both phonologically and phonetically. For example, in American English, **t** tends to be flapped (i.e. becomes '**d**'-like) in *capital*, but not in *military*. This is because flapping does not occur preceding stressed syllables (and *-tary* is stressed in American pronunciation). However, this difference is also carried over to other members of the paradigm, where the syllable stress is no longer different: **t** is flapped in *capitalistic* but not *militaristic*. The precise details of the phonetic

implementation of affixed words is affected by the other words present in the paradigm [67].

Another example concerns the devoicing of final obstruents in Dutch. In their infinitive forms the two Dutch verbs *verwijden* and *verwijten* differ in the identity of the medial consonant (**d** versus **t**). However, when the *-en* suffix is omitted, both words are produced with a **t**: *verwijt*. This is traditionally understood as a process which turns a discrete voiced /**d**/ into its discrete voiceless counterpart /**t**/. However, experiments have revealed that the process of devoicing is incomplete - the *verwijt* which relates to *verwijden* actually contains some **d**-like characteristics, which, furthermore, are functional for the listener [68].

The gradient morphological structure which emerges from paradigms is also reflected in details of sound structure. For example, words such as *swiftly* (which is more frequent than *swift*, and so only moderately segmentable) are pronounced with less of a **t** sound than matched words such as *softly* (less frequent than *soft*, and so highly segmentable) [14]. That is, words containing less support for their affixedness are associated with more phonetic reduction at the morphological boundary. An example from phonology comes from the use of intrusive **r** in New Zealand English: the insertion of **r** between base and suffix, as in *draw-r-ing*. This process is sensitive to the gradience of morphological structure, with affixes that receive more paradigmatic support more likely to attract **r** [69].

Gradeness and (ir)regularity

One of the central claims of the 'entries plus rules' model is that English learners would get by with a single default rule for regular verbs. However, there appear to be 'islands of reliability' within regular verbs which affect the probability of the regular past tense being used for novel verbs [70].

Another central claim of the entries plus rules model is that rules of inflection (such as the past tense) are sensitive only to the properties of a word's form, and blind to a word's meaning. After all, words like *shrink* and *drink* have the same kind of irregular past tense (*shrank*, *drank*), yet differ in meaning. And words with similar meanings like *slap* and *strike* can have very different past tense forms (*slapped*, *struck*). Recent studies [33,71], however, show that when irregular verbs are compared with regular verbs, irregular verbs turn out to have more semantic neighbors, and these semantic neighbors themselves are more likely to be irregular. The greater semantic density of irregulars is reflected in association norms, familiarity ratings, and chronometric measures of lexical processing, and points to a potential confound of (ir)regularity and semantic density in the brain imaging literature on regular and irregular verbs.

Conclusions

Is morphological structure inherently graded? The issue is controversial, but the evidence that is currently accumulating in the literature suggests that the answer is yes. Advances in theoretical morphology have led many scholars to reject the morpheme as a unit of analysis. Advances in statistics, data mining, and computational morphology have made it possible to develop formal models for paradigmatic analogy. Advances in linguistic data analysis have led to improved insight into the relation between (ir)regularity and semantics, and into the relation between graded structure on the one hand, and productivity, affix order, and phonetic realization on the other. Advances in auditory word recognition have documented listeners' sensitivity to fine phonetic detail in the acoustic form of simple and complex words, and its role in biasing the listener towards the correct meaning [49,72]. In short, recent developments suggest that the true complexity and, we would say, the true beauty of morphological structure can only be appreciated in full from a probabilistic perspective.

Accepting gradeness as part and parcel of the grammar entails a paradigm shift for linguistics. Especially generative linguistics has known a long history of antagonism with respect to the role of probability in the grammar. But the graded nature of morphological structure also challenges experimental approaches to lexical processing to develop models that are predictive not only for factorial extremes, but for the full range of intermediate cases as well.

The conclusion that morphological structure is indeed inherently graded does not necessarily imply that a subsymbolic, connectionist approach is called for. What makes the work of Rumelhart and McClelland and subsequent connectionist modeling exciting from a morphologist's point of view is that these are the first rigorous mathematical models for gradience in morphology. But artificial neural networks are but one of many currently available statistical tools for coming to grips with gradient structure [70,73–75]. The challenge for future research (see also Box 1) is to develop biologically plausible inductive models that do full justice to the – graded – structural intricacies of morphological complexity.

Box 1. Questions for further research

• How should analogical similarity be constrained, and what factors drive these constraints?
• What empirical evidence can establish whether stems and formatives develop independent representations?
• Most studies addressing morphological processing and representation have scrutinized words in isolation. How does context affect the balance of memory and exemplar-driven computation?
• To what extent does the evidence for gradeness generalize beyond experimental tasks?
• What empirical evidence can decide between subsymbolic generalization in neural networks and exemplar-driven symbolic generalization?
• How can semantics be realistically modelled in analogical models?
• To what extent can relative frequency and junctural phonotactics explain affix ordering in other languages?
• How do paradigms form between multiply-affixed words?
• If structure emerges gradiently from paradigms of encountered words, then an individual's social networks and personal linguistic experience could considerably affect the degree to which they represent and process specific words as morphologically complex. To what extent is this the case?

References

1 Pinker, S. and Prince, A. (1988) On language and connectionism. *Cognition* 28, 73–193

2 Pinker, S. (1997) Words and rules in the human brain. *Nature* 387, 547–548

3 Pinker, S. (1999) *Words and Rules: The Ingredients of Language*, Weidenfeld & Nicolson

4 Pinker, S. and Ullman, M. (2002) The past and future of the past tense. *Trends Cogn. Sci.* 6, 456–463

5 Pinker, S. and Ullman, M. (2002) Combination and structure, not gradedness, is the issue: Reply to McClelland and Patterson. *Trends Cogn. Sci.* 6, 472–474

6 Rumelhart, D.E. and McClelland, J.L. (1986) On learning the past tenses of English verbs. In *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*. (Vol. 2) *Psychological and Biological Models* (McClelland, J.L. and Rumelhart, D.E., eds), pp. 216–271, MIT Press

7 MacWhinney, B. and Leinbach, J. (1991) Implementations are not conceptualizations: revising the verb learning model. *Cognition* 40, 121–157

8 Plunkett, K. and Juola, P. (1999) A connectionist model of English past tense and plural morphology. *Cogn. Sci.* 23, 463–490

9 Seidenberg, M.S. and Gonnerman, L.M. (2000) Explaining derivational morphology as the convergence of codes. *Trends Cogn. Sci.* 4, 353–361

10 Joanisse, M.F. and Seidenberg, M.S. (1999) Impairments in verb morphology after brain injury: a connectionist model. *Proc. Natl. Acad. Sci. U. S. A.* 96, 7592–7597

11 McClelland, J.L. and Patterson, K. (2002) 'Words *or* Rules' cannot exploit the regularity in exceptions: Reply to Pinker and Ullman. *Trends Cogn. Sci.* 6, 464–465

12 Wurm, L.H. (1997) Auditory processing of prefixed English words is both continuous and decompositional. *J. Mem. Lang.* 37, 438–461

13 Hay, J. (2001) Lexical frequency in morphology: Is everything relative? *Linguistics* 39, 1041–1070

14 Hay, J. (2003) *Causes and Consequences of Word Structure*, Routledge

15 Gonnerman, L.M. and Anderson, E.S. (2001) Graded semantic and phonological similarity effects in morphologically complex words. In *Morphology 2000: Selected Papers from the 9th Morphology Meeting* (Bendjaballah, S. *et al.*, eds), pp. 137–148, Benjamins

16 Pastizzo, M.J. and Feldman, L.B. (2002) Discrepancies between orthographic and unrelated baselines in masked priming undermine a decompositional account of morphological facilitation. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 244–249

17 Plaut, D.C. and Gonnerman, L.M. (2000) Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Lang. Cogn. Process.* 15, 445–485

18 Rueckl, J.G. *et al.* (1997) Morphological priming, fragment completion, and connectionist networks. *J. Mem. Lang.* 36, 382–405

19 Beard, R. (1995) *Lexeme-morpheme Base Morphology: A General Theory of Inflection and Word Formation*, State University of New York Press

20 Aronoff, M. (1994) *Morphology by Itself: Stems and Inflectional Classes*, MIT Press

21 Anderson, S.R. (1992) *A-morphous Morphology*, Cambridge University Press

22 Erelt, M., ed. (2003) *Estonian Language*, Estonian Academy Publishers

23 Blevins, J.P. (2003) Stems and paradigms. *Language* 79, 737–767

24 Jackendoff, R.S. (1975) Morphological and semantic regularities in the lexicon. *Language* 51, 639–671

25 Sereno, J. and Jongman, A. (1997) Processing of English inflectional morphology. *Mem. Cogn.* 25, 425–437

26 Bertram, R. *et al.* (2000) Affixal homonymy triggers full-form storage even with inflected words, even in a morphologically rich language. *Cognition* 74, B13–B25

27 Sonnenstuhl, I. and Huth, A. (2002) Processing and representation of German -n plurals: a dual mechanism approach. *Brain Lang.* 81, 276–290

28 Baayen, R.H. *et al.* (2002) Dutch inflection: the rules that prove the exception. In *Storage and Computation in the Language Faculty* (Nooteboom, S. *et al.*, eds), pp. 61–92, Kluwer

29 New, B. *et al.* (2004) The processing of singular and plural nouns in French and English. *J. Mem. Lang.* 51, 568–585

30 Landauer, T.K. (1986) How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cogn. Sci.* 10, 477–493

31 Kostic, A. *et al.* (2003) Inflectional morphology and word meaning: orthogonal or co-implicative domains? In *Morphological Structure in Language Processing* (Baayen, R.H. and Schreuder, R., eds), pp. 1–44, Mouton de Gruyter

32 del Prado Martin, F.M. *et al.* (2004) Putting the bits together: An information theoretical perspective on morphological processing. *Cognition* 94, 1–18

33 Tabak, W. *et al.* Lexical statistics and lexical processing: semantic density, information complexity, sex, and irregularity in Dutch. In *Linguistic Evidence – Empirical, Theoretical, and Computational Perspectives* (Kepser, S. and Reis, M., eds), Mouton de Gruyter (in press)

34 Moscoso del Prado Martin, F. *et al.* (2004) Morphological family size in a morphologically rich language: The case of Finnish compared to Dutch and Hebrew. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 1271–1278

35 Ernestus, M. and Baayen, R.H. (2003) Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79, 5–38

36 Ernestus, M. and Baayen, R.H. (2004) Analogical effects in regular past tense production in Dutch. *Linguistics* 42, 873–903

37 Gagne, C. and Shoben, E.J. (1997) The influence of thematic relations on the comprehension of modifier-noun combinations. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 71–87

38 Krott, A. *et al.* (2001) Analogy in morphology: modeling the choice of linking morphemes in Dutch. *Linguistics* 39, 51–93

39 Krott, A. *et al.* (2004) Sublexical units and supralexical combinatorics in the processing of interfixed Dutch compounds. *Lang. Cogn. Process.* 19, 453–471

40 Sandra, D. and Fayol, M. (2003) Spelling errors as another way into the mental lexicon. In *Morphological Structure in Language Processing* (Baayen, R.H. and Schreuder, R., eds), pp. 485–514, Mouton de Gruyter

41 Reid, A.A. and Marslen-Wilson, W.D. (2003) Lexical representation of morphologically complex words: Evidence from Polish. In *Morphological Structure in Language Processing* (Baayen, R.H. and Schreuder, R., eds), pp. 287–336, Mouton de Gruyter

42 Burani, C. and Thornton, A.M. (2003) The interplay of root, suffix and whole-word frequency in processing derived words. In *Morphological Structure in Language Processing* (Baayen, R.H. and Schreuder, R., eds), pp. 157–208, Mouton de Gruyter

43 Taft, M. (1981) Prefix stripping revisited. *J. Verbal Learn. Verbal Behav.* 20, 289–297

44 Seidenberg, M. (1987) Sublexical structures in visual word recognition: Access units or orthographic redundancy. In *Attention and Performance XII* (Coltheart, M., ed.), pp. 245–264, Erlbaum

45 Laudanna, A. and Burani, C. (1995) Distributional properties of derivational affixes: Implications for processing. In *Morphological Aspects of Language Processing* (Feldman, L.F., ed.), pp. 345–364, Erlbaum

46 Bergen, B.K. (2004) The psychological reality of phonaesthemes. *Language* 80, 290–311

47 Feldman, L.B. (2000) Are morphological effects distinguishable from the effects of shared meaning and shared form? *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1431–1444

48 Navon, D. (1977) Forest before trees: the precedence of global features in visual perception. *Cogn. Psychol.* 9, 353–383

49 Salverda, A.P. *et al.* (2003) The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition* 90, 51–89

50 Schreuder, R. *et al.* (2003) Parsing and semantic opacity. In *Reading Complex Words. Cross-Language Studies* (Assink, E.M.H. and Sandra, D., eds), pp. 159–189, Kluwer

51 Bowers, J.S. *et al.* (2005) Automatic semantic activation of embedded words: Is there a 'hat' in 'that'? *J. Mem. Lang.* 52, 131–143

52 Taft, M. (1994) Interactive-activation as a framework for understanding morphological processing. *Lang. Cogn. Process.* 9, 271–294

53 Longtin, C. *et al.* (2003) Morphological priming without morphological relationship. *Lang. Cogn. Process.* 18, 313–334

54 Taft, M. (2004) Morphological decomposition and the reverse base frequency effect. *Q. J. Exp. Psychol. A.* 57, 745–765

55 Baayen, R.H. (2003) Probabilistic approaches to morphology. In *Probabilistic Linguistics* (Bod, R. *et al*., eds), pp. 229–287, MIT Press

56 Bauer, L. (2001) *Morphological Productivity*, Cambridge University Press

57 Baayen, R.H. and Renouf, A. (1996) Chronicling The Times: productive lexical innovations in an English newspaper. *Language* 72, 69–96

58 Hay, J.B. and Baayen, R.H. (2003) Phonotactics, parsing and productivity. *Ital. J. Linguistics* 1, 99–130

59 Booij, G.E. (2001) Lexical phonology: a review. In *Phonology: Critical Concepts* (Kreidler, C., ed.), pp. 203–235, Routledge

60 Fabb, N. (1988) English suffixation is constrained only by selectional restrictions. *Natural Lang. Linguist. Theory* 6, 527–539

61 Plag, I. (1996) Selectional restrictions in English suffixation revisited: a reply to Fabb (1988). *Linguistics* 34, 769–798

62 Giegerich, H.J. (1999) *Lexical Strata in English: Morphological Causes, Phonological Effects*, Cambridge University Press

63 Benua, L. (2000) *Phonological Relationships Between Words*, Garland Publishing

64 Hay, J.B. (2002) From speech perception to morphology: Affix-ordering revisited. *Language* 78, 527–555

65 Bybee, J.L. (1985) *Morphology: A study of the Relation Between Meaning and Form*, Benjamins

66 Hay, J.B. and Plag, I. (2004) What constrains possible suffix combinations? On the interaction of grammatical and processing restrictions in derivational morphology. *Natural Lang. Linguist. Theory* 22, 565–596

67 Steriade, D. (2000) Paradigm uniformity and the phonetics-phonology boundary. In *Papers in Laboratory Phonology 5* (Broe, M. and Pierrehumbert, J., eds), pp. 313–335, Cambridge University Press

68 Ernestus, M. and Baayen, R.H. (in press) The functionality of incomplete neutralization in Dutch: The case of past-tense formation. In *Laboratory Phonology 8* (Goldstein, L.M. *et al.*, eds), Mouton de Gruyter

69 Hay, J.B. and Warren, P. (2003) Experiments on /r/-intrusion. *Wellington Working Papers in Linguistics* 14, 47–58

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

71 Baayen, R.H. and Moscoso del Prado M.F. Semantic density and past-tense formation in three Germanic languages. *Language* (in press)

72 Kemps, R. *et al.* (2005) Prosodic cues for morphological complexity in Dutch and English. *Lang. Cogn. Process.* 20, 43–73

73 Boersma, P. and Hayes, B. (2001) Empirical tests of the gradual learning algorithm. *Linguist. Inquiry* 32, 45–86

74 Skousen, R. (2002) *Analogical Modeling*, Benjamins

75 Daelemans, W. *et al.* (2001) *TiMBL: Tilburg Memory Based Learner Reference Guide. Version 4.2*. Technical Report ILK 02-01, Computational Linguistics Tilburg University