In S. Bendjaballah, W. U. Dressler, O. E. Pfeiffer & M. D. Voeikova (eds.), Morphology 2000, pp. 213-225. Amsterdam: John Benjamins, 2002.

CHAPTER 17

Is there a morphological parser?

Gary Libben and Roberto G. de Almeida
University of Alberta and Concordia University

In the psycholinguistic literature on lexical processing, the term "morphological parser" is most often used to refer to the hypothesized component of the human language processing system that is responsible for the isolation and identification of morphological constituents of multimorphemic words. This morphological parser makes it possible for language users to understand novel or infrequent multimorphemic words by breaking such words into their morphological constituents so that an interpretation of the novel form can be constructed on the basis of those constituents.

While it is relatively uncontroversial to assume that morphological parsing plays a key role in the processing of novel multimorphemic words, its status in the recognition of existing multimorphemic words is much less clear. Over the past quarter century, two opposing hypotheses about the role of morphological parsing in visual word recognition have dominated the lexical processing landscape. On the one hand, the morphological decomposition hypothesis (see, for example, Taft and Forster 1975) has claimed that complex word forms such as *unbelievable* gain access to the lexicon by a procedure that first extracts the word's affixes (*un-* and *-able*) and then uses the word's stem (*believe*) to find an entry in the mental lexicon. Under this hypothesis, parsing is claimed to be a pre-lexical morphological operation in that it occurs prior to the recognition of the whole word. The full-listing hypothesis, on the other hand (Butterworth 1983, for instance), has assumed that full forms are the basis for word recognition and that, if morphological analysis takes place at all, it occurs after a word has gained access to the lexicon. Thus, under this view, morphological analysis is a post-access process rather than a pre-lexical parsing procedure.

The two hypotheses make distinct claims concerning the nature of lexical representation codes. For the decomposition hypothesis, lexical entries are roots or stems, while for the full-listing hypothesis, most morphologically complex forms have independent entries in the mental lexicon. Thus far, both hypotheses have found some support. This has been a primary motivation in the elaboration of dual route models (see Baayen, Dijkstra and Schreuder 1997; Laudanna and Burani 1995; Schreuder and Baayen 1995) in which morphological decomposition procedures and whole word processing procedures yield separate and competing analyses. Such elaborations have allowed for a more adequate treatment of the diversity of morphological priming effects. However, the question of whether the primary locus of morphological effects in visual word recognition is pre- or post-access has remained unresolved (McQueen and Cutler 1998).

In this chapter, we take up the question of whether it is possible to determine the extent to which pre-lexical morphological parsing takes place in the recognition of existing words of English. We assume as a background to this investigation a dual route perspective in which a native speaker of English is in possession of a morphological parser that is certainly employed when novel or infrequent multimorphemic words are encountered (see Libben, Derwing and de Almeida 1999). We further assume that some form of post-lexical morphological analysis also takes place in the recognition of existing complex and compound words. This post-lexical analysis may result from connections among lexical representations or from structured information within representation (see Libben in press, for a discussion of these alternatives).

The two assumptions above are critical to the conceptual structure of our report. First, the assumption of routine morphological parsing for novel words suggests that pre-lexical morphological parsing is in principle available to the recognition of real words. Second, the assumption of post-lexical morphological processing for existing words suggests that morphological parsing is not necessary for the recognition of those words. Thus, our question becomes: Is morphological parsing employed even in cases where it is not required?

In our view, framing the question in this manner has a critical advantage in allowing us to tease apart pre-lexical and post-lexical morphological operations. A long-standing problem in this enterprise has been the equivocal nature of data from priming studies that yield evidence of constituent activation in multimorphemic word recognition (see Sandra 1994; McQueen and Cutler 1998). The essence of the problem is that both pre-lexical morphological parsing and post-lexical processes should yield the same results for most multimorphemic words. Both processes, for example, would generate activation of *black* and *board* in the recognition of *blackboard*. Because most psycholinguistic techniques measure the results of lexical activation after it has occurred, determining the sequence of whole-word and constituent activation is not an easy matter.

How then might it be possible to distinguish between pre-lexical and post-lexical morphological operations? One means would be to seek cases in which the two processes do not yield the same results, so that the information provided by each operation is in a sense 'tagged' with respect to its source.

Below, we present two sources of evidence that move us toward the satisfaction of this constraint. We first summarize a recent investigation of aphasic interpretation errors in the reading of semantically opaque compounds that points to the view that pre-lexical and post-lexical morphological analyses contribute independently to lexical comprehension. We follow this with the presentation of a new experimental technique that was designed to measure these independent contributions. The results that we report using this technique support the view that pre-lexical morphological parsing is an obligatory and automatic component of the lexical processing system. Thus, we conclude: "There is a morphological parser and it is always on". In the final section of this chapter, we discuss the consequences of this conclusion for our conceptions of the characteristics of morphological parsing and its place in the lexical processing system.

1. Semantically opaque compounds

Semantically opaque compounds have been the subject of attention in the psycholinguistic literature because evidence of decomposition for these forms points directly to the role of pre-lexical morphological parsing in multi-morphemic word recognition (Jarema, Busson, Nikolova, Tsapkini and Libben 1999; Sandra 1990; Zwitserlood 1994). As we have pointed out above, both pre-lexical and post-lexical morphological processing might yield activation of the constituents *lady* and *bug* in the compound *ladybug*. However, there seems little reason to expect that post-lexical operations would yield activation of the words *hum* and *bug* for the semantically opaque form *humbug*. Any such activation, therefore, would be due to the effects of a pre-lexical parser.

If, indeed, a pre-lexical morphological parser yields constituent activation for semantically opaque forms, this activation would need to be discarded or suppressed soon after whole-word activation. The reason for this is that the information supplied by pre-lexical morphological parsing and post-lexical morphological analysis is semantically incompatible.

Libben (1993, 1998) described the case of an aphasic patient, RS, whose reading and interpretation of semantically opaque compounds indicated the preservation of both sources of lexical activation. When asked, for example, what the word *butterfly* means, she responded "a pretty yellow fly". Note here that she was blending together both the meaning of the whole word and the meaning of the constituents. The word "pretty" in her response was assumed to be an associate of the whole word *butterfly* (because most flies are not perceived to be pretty). In contrast, the word "yellow" was assumed to be an associate of the constituent *butter*. It seems unlikely that this activation would result from post-lexical processing. Rather, Libben (1998) argued that it is the consequence of a morphological parse of the word, the results of which are not discarded or suppressed during interpretation. This conclusion is supported by similar interpretations that RS showed for other opaque words. For *summersault* she said "you roll on the grass in the summer" again blending together whole word and constituent meaning. Finally, when asked the meaning of the word *dumbbell*, she responded "stupid weights . . . Arnold". In this case she showed activation of the whole word meaning of *dumbbell*, which is a type of exercise weight, the constituent *dumb*, and the association to Arnold Schwarzenegger, the former bodybuilder.

Although RS's pattern of compound reading is unusual, it is apparently not unique. McEwen, Westbury, Buchanan and Libben (in press) describe a deep dyslexic, JO, who showed a similar error pattern or performance with these words. Her reading for *butterfly* was "bread, butter, ..fly". The word "bread" could again only be an associate of the constituent *butter* and not of the compound *butterfly*. Similarly when attempting to read the semantically opaque compound *pancake*, JO produced the associates "breakfast" and "birthday". The former is related to the meaning of the whole word, while the second is related only to the constituent *cake*.

In our view, error patterns such as these provide intriguing evidence for a dissociation of information provided by the morphological parser and information provided through the representation of the multimorphemic string in the mental lexicon. We use the term intriguing rather than conclusive, however, because there are alternative accounts of these data. Although compounds such as *butterfly* and *pancake* are described as semantically opaque, it is nevertheless possible that the activation of their constituents results from the nature of the representations of the compounds in the mental lexicon rather than from a pre-lexical morphological parser (see, for example, Zwitserlood 1994).

We return, therefore, to the desideratum expressed above. The most revealing stimuli in the study of pre-lexical morphological parsing for existing words are most likely those for which whole-word morphological analysis and pre-lexical parsing yield different information. In the section below, we report on an initial study in which we have attempted to isolate such stimuli and to investigate their processing patterns in a semantic priming experiment.

2. Suffixed ambiguous roots

In the experiment described below, we relied on findings in the lexical-semantic access literature — in particular, in the lexical ambiguity resolution literature — to investigate the nature and locus of the morphological parsing process. It is well known that when participants encounter an ambiguous word such as *bark* in isolation or in a neutral context, both of its meanings — the one associated with *tree* and the one associated with *dog* — are accessed in the early stages of lexical processing (see Onifer and Swinney 1981; Seidenberg, Tanenhaus, Leiman and Bienkowski 1982). Studies of lexical ambiguity also find effects of meaning dominance, with the meaning that is used more frequently yielding higher priming effects than the less frequent meaning, both in isolation and in sentential contexts (Hogaboam and Perfetti 1975; Holmes 1979; Tabossi, Colombo and Job 1987). In this study, we used these findings as the basis of an investigation of whether affixed words such as *barking* — composed of an ambiguous root such as *bark* and a suffix such as *-ing* — are decomposed prior to lexical access. We reasoned that if pre-lexical morphological parsing occurs, *barking* would prime both *tree* and *dog*, because both meanings of the root *bark* would be accessed. On the other hand, if morphological parsing does not occur (that is, if the entire lexical item is used for access), then a word such as *barking* should prime only *dog*, because the *-ing* suffix serves to disambiguate the root before the word's semantic representation is accessed. Thus, although the word *bark* in isolation is ambiguous, the suffixed form *barking* is not, because its internal representation ([[bark]$_V$ing]) can only contain the verb *bark*. This verb representation may be associated with the meaning *dog* but not with the meaning *tree*.

2.1 Participants

Thirty undergraduate students from the University of Alberta participated as volunteers in this experiment. They were all native speakers of English attending introductory linguistic courses and were not informed about the purposes of the experiment.

2.2 Materials and Design

Twenty-four suffixed ambiguous roots served as the core stimuli for this experiment. The selected roots had the following characteristics: they were all monomorphemic words that could either be understood as nouns or as semantically unrelated verbs (for example *bark*). The addition of the suffix *-ing* to these words selects only the verb reading and, thus, disambiguates the root both syntactically and semantically.

The uninflected form of each word had a dominant and a subordinate meaning. Meaning dominance was determined by an association task with 137 native speakers of English, all attending an introductory Psychology course at Rutgers University. In this task, participants were presented with a list of 106 ambiguous words and, for each one, they were asked to write down the first word that came to mind. Dominant meanings were selected based on the relative frequency with which words related to each meaning were given by the participants. This association task allowed us to choose 24 core stimuli which had two distinct and unrelated meanings — that is, they were homonyms rather than polysemes.

The core ambiguous stimuli were organized with their semantic associates into prime-target pairs. In each pair, the suffixed ambiguous root served as the prime and its associates (as determined by the association task described above) served as the targets. Thus each ambiguous root resulted in the creation of two prime-target pairs (*barking-dog*, *barking-tree*, for instance). These target pairs were matched to control stimulus pairs that consisted of frequency matched primes and identical targets (such as *buck-dog* and *buck-tree*). Of the 24 ambiguous roots, 12 selected the dominant meaning and 12 selected the subordinate meaning when affixed by *-ing*. So, for instance, while *barking* is related to *dog*, an associate of the dominant meaning of *bark*, *training* is related to *fitness*, an associate of the subordinate meaning of *train*. We reasoned that if pre-lexical decomposition occurs and is sensitive to meaning dominance, *training* should prime *station* (an associate of the dominant meaning of *train*) as much as *barking* primes *dog*. This manipulation was necessary also because if priming effects were not obtained between *barking* and *tree*, this could be due to the fact that *tree* is related to the subordinate meaning of *bark*. But in the case of *training-station*, priming effects can

Table 1. Prime and target pairs

Primes	Targets		Primes	Targets	
	Dom	Sub		Dom	Sub
banking	*teller*	river	interesting	money	*bore*
barking	*dog*	tree	jerking	stupid	*pull*
faning	*air*	sports	leaning	fat	*stand*
gagging	*choke*	joke	littering	cat	*garbage*
majoring	*minor*	army	lobbying	hotel	*congress*
partying	*fun*	political	matching	fire	*mix*
perching	*bird*	fish	ringing	finger	*phone*
pitching	*ball*	sound	sentencing	word	*judge*
slipping	*fall*	paper	shedding	garage	*hair*
spelling	*write*	witch	springing	summer	*jump*
stapling	*gun*	crop	squashing	vegetable	*flat*

Note: There are two targets for each prime. The first is a semantic associate of the dominant meaning. The second is a semantic associate of the subordinate meaning. In the left half of the table, the addition of the *-ing* affix biases interpretation toward the dominant meaning of the root (italicized). In the right half of the table, the addition of *-ing* biases toward the subordinate meaning (italicized).

only be attributed to pre-lexical decomposition. The entire set of core stimuli is presented in Table 1.

2.3 Procedure

We employed a visual masked priming lexical decision technique. Subjects saw a sequence of four events on the screen: (1) they saw a fixation point (an asterisk) for 1900 milliseconds (ms), (2) a mask composed by row of hash marks ("#") presented for 500 ms, (3) the prime presented in lowercase letters for 80 ms, and (4) the target presented in uppercase letters for 500 ms. Subjects were instructed to make a lexical decision (word or nonword) on the string of capital letters by pressing either a button labeled *yes* if the string formed an English word, or a button labeled *no* otherwise. They were not told that lowercase primes preceded the targets, but if they noticed the presence of the primes during the practice trials, they were instructed to pay attention only to the targets. A new trial began — with the fixation point — as soon as the subject pressed one of the buttons. Responses were timed from the onset of the target until the subject pressed a button. The stimuli were all presented sequentially on the middle of the screen, in white Courier New 24 font over black background. The experiment was run on Macintosh PowerBook 520 computers running PsyScope (Cohen, MacWhinney, Flatt and Provost 1993) equipped with external Apple monochromatic monitors and PsyScope button boxes.

2.4 Results and Conclusions

Prior to statistical analysis, all response latencies less than 300 milliseconds (ms) and greater than 1,200 ms (2 percent of the responses) were eliminated from the data set. The data were first analyzed in a 2 × 2 × 2 Analysis of Variance with repeated measures on all factors (prime type, association dominance, and affixation bias). The means and standard deviations for this analysis are provided in Table 2.

As can be seen in Table 2, there is no overall priming effect ($F(1,29)=.33$, $p=.56$). The absence of this priming effect results from an interaction in the data between the other two factors, 'association dominance' and 'affixation bias'. The pattern of this interaction may be described as follows: When there is consistency between the affixed prime and the target, a small priming effect of 16 ms is observed. These are cases such as *barking* priming *dog* and *ringing* priming *phone*. The opposite pattern occurs for cases in which the associate is inconsistent with the affixed form of the ambiguous root (for example, *barking-tree* and *ringing-finger*). Here we see negative priming effects the magnitude of which are related to the meaning dominance of the ambiguous root. Thus, the greatest negative priming effect (–34 ms) is seen in the case in which the semantic associate that is disallowed as a result of suffixation is also the dominant associate of the ambiguous root (*ringing-finger*, for instance).

Table 2. Mean response latencies and standard deviations (SD) for all prime-target relations

Example stimuli		Control prime		Exper. prime		Priming effect
Prime	Target	RT	SD	RT	SD	
Barking	*dog*	583	(86)	567	(104)	16
	tree	572	(93)	591	(125)	–19
Ringing	finger	549	(69)	583	(102)	–34
	phone	595	(98)	579	(95)	16
Average		574.7		580		–5

Note: The first two columns represent example stimuli for each condition. The final column shows the priming effect as the RT of the control prime condition minus the RT of the experimental prime condition.

Thus the results of this experiment show a significant effect of the consistency between the effects of suffixation and semantic association to the ambiguous root. This interaction effect is presented graphically in terms of priming effects in Figure 1. As can be seen in this figure, the data reveal no significant main effects of association dominance ($F(1,29)=.227$, $p=.64$) or affixation bias ($F(1,29)=.316$, $p=.58$), but a significant interaction between these two factors ($F(1,29)=15.4$, $p=.0005$).

In our view, these results point very strongly to the view that pre-lexical morphological parsing occurs for these relatively common suffixed words. This conclusion is based on the negative priming effect that was obtained for inconsistent prime-target pairs. If no pre-lexical parse of these stimuli occurred during the priming phase of each experiment trial, the inconsistent primes should have behaved exactly as their unrelated control words. The fact however, that they were significantly slower than their controls suggest that at some point during lexical processing, both meanings of the ambiguous roots were activated. As we have argued above, this double activation could only result from

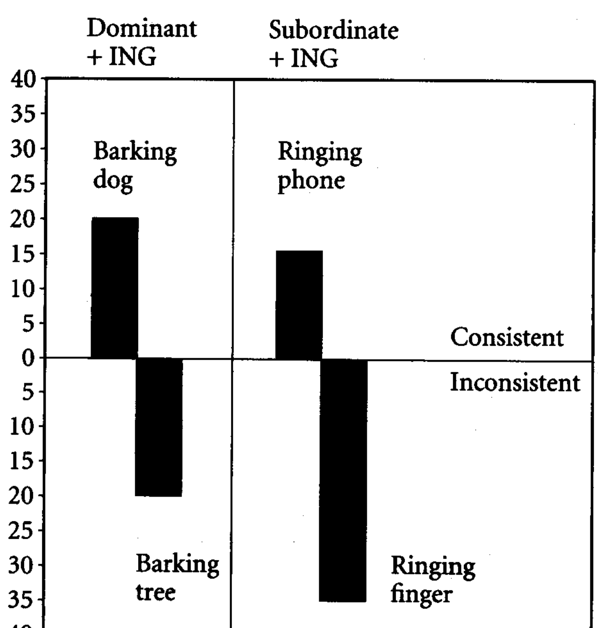

Figure 1. Response latency data viewed as priming effects (control primed condition minus experimental prime condition)

pre-lexical morphological parsing. Thus for prime-target pairs such as *barking-tree* and *ringing-finger*, participants received inconsistent information. This inconsistency, we reason, resulted in hesitation and the elevated response times observed in the data.

3. General Discussion

Our goal at the outset of this study was to address the question of whether pre-lexical morphological parsing occurs as a component of normal lexical processing. As we have noted, there have been significant methodological barriers to the treatment of this question because constituent activation — the normal result of pre-lexical parsing — can also occur as a result of post-access processes in the mental lexicon. In this chapter, we have argued that this difficulty can be overcome by isolating stimulus categories for which the information yielded by pre-lexical parsing is not identical to that which would be expected to obtain from post-lexical operations.

Semantically opaque compounds may be regarded as one such stimulus category. In our presentation of RS's interpretation of semantically opaque compounds, we argued that lexical processing impairment in aphasia can result from the inability to discard or suppress information arising from pre-lexical parsing when this information is inconsistent with information obtained from whole-word access. This line of reasoning suggests that for non-

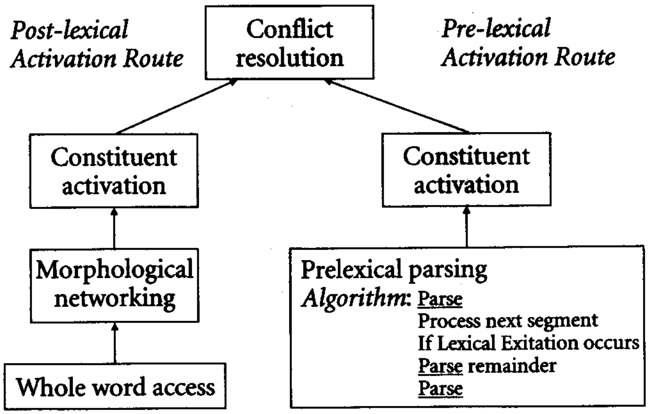

Figure 2. The interaction of pre-lexical parsing and whole-word access in visual word recognition

aphasic native speakers of English, these two sources of information also lead to inconsistencies and activation conflict that must be resolved. If indeed, pre-lexical morphological parsing plays an integral role in normal lexical processing, the process of conflict resolution would also be expected to play a key role in the lexical system's functional architecture. This claim is represented graphically in Figure 2.

The organization of the model in Figure 2 highlights the role of 'conflict resolution' against the background of a dual-route model. We assume that morphological parsing is an automatic and obligatory component of the word recognition process for both existing and novel multimorphemic forms. We expect that morphological parsing (shown on the right side of the figure) interacts with a generally faster whole-word recognition process (shown on the left side of the figure) that is also always "on". Both whole word access and pre-lexical decomposition typically result in constituent activation. However, in the case of whole word recognition, that constituent access derives from the representation for the complex word. It therefore always gets the right constituents. Pre-lexical decomposition, on the other hand, can both over-generate and mis-generate. This creates the need for conflict resolution (represented in the upper center of the model).[1]

Our interpretation of the data from the suffixed ambiguous root experiment is that the dominant result, the negative priming effect for inconsistent prime-target pairs, reflects a stage in processing before conflict resolution have been completed. The consequence of the unresolved conflict is an inhibition effect, rather than the priming effect we had originally predicted, or the null effect, which would be predicted by the full-listing hypothesis.

Finally, if this interpretation is correct and the representation in Figure 2 appropriately depicts the manner in which pre-lexical and post-lexical analysis interact, we are led to the prediction that the inhibition effect we observed with an 80 ms prime may be expected to dissipate with longer prime durations and prime-target intervals. Thus, in our view, further experimentation with stimuli such as these would have a promising role to play in increasing our understanding of the time-course of conflict resolution and ultimately our understanding of the functional architecture of morphological processing.

Note

1. The pre-lexical parsing algorithm provided in the model is based on the proposal in

Libben (1994) and essentially claims that pre-lexical morphological parsing can be described as a recursive procedure that isolates morphemes in a beginning-to-end manner across an input string (see also Andrews and Davis 1999; Hudson and Buis 1995). The role of morphological networking in post-lexical activation has been discussed in Libben (in press) where it is claimed that morphological constituency effects result from the architecture of the morphological networks for individual words and the manner in which components of those networks are activated.

References

Andrews, S. and Davis, C. 1999. "Interactive activation accounts of morphological decomposition: finding the trap in mousetrap". *Brain and Language* 68.355–61.

Baayen, R.H., Dijkstra, T. and Scheuder, R. 1997. "Singulars and plurals in Dutch: Evidence for a parallel dual-route model". *Journal of Memory and Language* 37.94–117.

Butterworth, B. 1983. "Lexical representation." *Language Production 2.* ed. by B. Butterworth (257–94). New York, Academic Press.

Cohen, J., MacWhinney, B., Flatt, M. and Provost, J. 1993. "PsyScope: An interactive graphical system for designing and controlling experiments in the psychology laboratory using Macintosh computers". *Behavior Methods, Research, Instruments, and Computers* 25.257–71.

Hogaboam, T.W. and Perfetti, C.A. 1975. "Lexical ambiguity and sentence comprehension". *Journal of Verbal Learning and Verbal Behavior* 14.265–74.

Holmes, V.M. 1979. "Accessing ambiguous words during sentence comprehension". *Quarterly Journal of Experimental Psychology* 31.569–89.

Hudson, P.T.W. and Buis, D. 1995. "Left-to-right processing of derivational morphology". *Morphological aspects of language processing.* ed. by L.B. Feldman (383–96). Hillsdale, NJ: Lawrence Erlbaum.

Jarema, G., Busson, C., Nikolova, R., Tsapkini, K. and Libben, G. 1999. "Processing compounds: A cross-linguistic study". *Brain and Language* 68.362–9.

Laudanna, A. and Burani, C. 1995. "Distributional properties of derivational affixes: Implications for processing." *Morphological aspects of language processing.* ed. by L.B. Feldman (345–64). Hillsdale, NJ: Lawrence Erlbaum.

Libben, G. 1993. "A case of obligatory access to morphological constituents". *The Nordic Journal of Linguistics* 16.111–21.

——1994. "How is morphological decomposition achieved?" *Language and Cognitive Processes* 9.369–91.

——1998. "Semantic transparency in the processing of compounds: consequences for representation, processing, and impairment". *Brain and Language* 61.30–44.

——in press. "Morphological Parsing and Morphological Structure". *Reading complex words,* ed. by A. Egbert and D. Sandra. Amsterdam: Kluwer.

——Derwing B. and de Almeida, R.G. 1999. "Ambiguous novel compounds and models of morphological parsing". *Brain and Language* 68.378–86.

McEwen, S., Westbury, C., Buchanan, L. and Libben, G. in press. "Semantic information is used by a deep dyslexic to parse compounds". *Brain and Cognition.*

McQueen, J. and Cutler, A. 1998. "Morphology in word recognition." In: A. Spencer and A. Zwicky (eds.), The handbook of morphology (406–27). Oxford: Blackwell.

Onifer, W. and Swinney, D.A. 1981. "Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias". *Memory and Cognition* 9.225–36.

Sandra, D. 1990. "On the representation and processing of compound words: automatic access to constituent morphemes does not occur". *Quarterly Journal of Experimental Psychology* 42.529–67.

——1994. "The morphology of the mental lexicon. Internal word structure viewed from a psycholinguistic perspective". *Language and Cognitive Processes* 9.227–69.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M. and Bienkowski, M. 1982. "Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge based processing". *Cognitive Psychology* 14.489–537.

Schreuder, R. and Baayen, R.H. 1995. "Modelling morphological processing." *Morphological aspects of language processing.* ed. by L.B. Feldman (131–56). Hillsdale, NJ: Lawrence Erlbaum.

Tabossi, P., Colombo, L. and Job, R. 1987. "Accessing lexical ambiguity: Effects of context and dominance". *Psychological Research* 49.161–7.

Taft, M. 1981. "Prefix stripping revisited." *Journal of Verbal Learning and Verbal Behavior* 20.289–97.

——and Forster, K.I. 1975. "Lexical storage and the retrieval of prefixed words." *Journal of Verbal Learning and Verbal Behavior* 14.630–47.

Zwitserlood, P. 1994. "The role of semantic transparency in the processing and representation of Dutch compounds". *Language and Cognitive Processes* 9.341–68.

This is an offprint from:

S. Bendjaballah, W.U. Dressler, O.E. Pfeiffer and M. Voeikova (eds)
Morphology 2000.
Selected papers from the 9th Morphology Meeting,
Vienna, 24-28 February 2000.
John Benjamins Publishing Company
Amsterdam/Philadelphia
2002
(Published as Vol. 218 of the series
CURRENT ISSUES IN LINGUISTIC THEORY,
ISSN 0304-0763)

ISBN 90 272 3725 5 (Hb; Eur.) / 1 58811 080 X (Hb; US)