Is there a morphological Parser?

Gary Libben, University of Alberta

In this paper, I address the question of how native speakers of a language access the morphological constituents of multimorphemic strings. This ability is fundamental to native speaker competence and is of course the only means by which novel words such as *unfillability* and *bookfloor* can be understood. Over the past quarter century, a good deal of research has investigated whether existing multimorphemic words are also decomposed into their constituents as part of the word recognition process (e.g., Sandra 1990, Henderson 1985) and the manner in which this decomposition is achieved (e.g., Taft and Forster 1975, Taft 1981, Libben 1994, Hudson & Buis 1995, Andrews and Davis 1999). This research has led to a rather complex view in which the interplay between morphological decomposition and whole-word recognition is influenced by factors such as frequency, semantic transparency, morphological structure, and morphological class (see McQueen & Cutler, 1998 for a review). We have thus learned a great deal about the role of stimulus variation in experiments that target morphological decomposition, yet fundamental questions concerning the nature of morphological decomposition itself remain unresolved. These include:

1. What mechanisms allow constituent morphemes to be identified?
2. Are the same procedures used in parsing prefixed, suffixed, and compound words?
3. How does morphological parsing interact with whole word recognition?
4. Is it possible to distinguish between pre-lexical and post-lexical decomposition?

In this paper, I present evidence from three sets of psycholinguistic studies that address these questions.

The first set of experiments deals with German interfixed compounds such as *Firmensitz, Koenigshof*, and *Farmenkauf*. We have found that in both the decomposition and composition of these compounds, the length and complexity of the interfix plays a significant role. This suggests that participants in the experiments parse these compounds into their constituents rather than simply "look them up" subsequent to whole word recognition.

The second set of experiments addressed the question of the mechanisms involved in morphological parsing. We employed a set of ambiguous novel compounds in English. These are words such as *clamprod* that can be parsed as either *clam-prod* or *clamp-rod*. Our results suggest that morphological parsing operates in a left-to-right manner but does not simply divide a compound into its constituents. Rather, it generates activation for all constituent morphemes (i.e., *clam, clamp, prod* and *rod*). Finally, our data suggest that multiple combinations of initial and final constituents are activated (i.e., *clamp-rod, clam-prod* and also possibly *clamp-prod*). Libben, Derwing and de Almeida (1999) argue that this pattern of results suggests that the operation of the morphological parser is not characterized by processing efficiency. Rather is appears to be designed to generate the maximum number of legal morphological structures within a word.

It is important to note, however, that the Libben, Derwing and de Almeida (1999) study dealt with novel compounds, and therefore the results do not address the question of whether existing words also show evidence of left-to-right prelexical morphological parsing. In our third set of experiments, we probed morphological

parsing by examining the processing of words such as *barking* and *interesting*. The key characteristic of these stimuli is that their stems are ambiguous but the affixed words are not. So, whereas *bark* may refer to the outermost layer of a tree, or the sound that a dog makes, the affixed form normally only has the latter meaning. Thus, if presentation of *barking* activates both the lexical associates *dog* and *tree*, we can assume that this activation must result from prelexical morphological decomposition.

Taken together, the results of these studies suggest that prelexical morphological decomposition is an automatic and obligatory component of the word recognition process for both existing and novel multimorphemic forms. I argue that the mechanisms by which constituent morphemes are identified can be described in terms of a serial processing algorithm that does not distinguish among prefixed, suffixed and compound words. Although morphological decomposition is an integral part of word recognition, it interacts with a generally faster whole-word recognition process that prevents, for example, the activation of *bar* and *king* in the word *barking*, but allows activation of *bark* and *–ing*.

So there is a morphological parser – but its character is different from the one we might have imagined at the outset. It is slow, inefficient, over-productive, and error-prone. But it is also always "on".

References

Andrew, S. & Davis, C. 1999. "Interactive activation accounts of morphological decomposition: finding the trap in mousetrap." Brain and Language, 68, 355-361.

Henderson, L. 1985. "Toward a psychology of morphemes." In: A. W. Ellis (eds.), Progress in the psychology of language, vol. 1, 15-72.

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

Libben, G. 1994. "How is morphological decomposition achieved?" Language and Cognitive Processes, 9, 369-391.

Libben, G., Derwing B., & de Almeida, R. G. 1999. "Ambiguous novel compounds and models of morphological parsing". Brain and Language, 68, 378-386.

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

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, 529-567.

Taft M. & Forster, K.I. 1975. "Lexical storage and the retrieval of prefixed words." Journal of Verbal Learning and Verbal Behavior, 14, 630-647.

Taft, M. 1981. "Prefix stripping revisited." Journal of Verbal Learning and Verbal Behavior, 20, 289-297.