Morphology, modality, and lexical architecture

WILLIAM MARSLEN-WILSON, XIAOLIN ZHOU AND MIKE FORD

1. INTRODUCTION

In thinking about how the human language system is organised to mediate the relationship between internally represented knowledge and the input and output systems dedicated to the access and use of this knowledge, it is natural to assume that the system has a considerable degree of functional and architectural symmetry. In current models of the organisation of the mental lexicon (e.g., Miceli 1994, Seidenberg 1995), we see diagrams very much like the one illustrated in Figure 1, with a central, modality-independent store of lexical content, and parallel sets of input lexica for the two principal modalities (speech and vision).
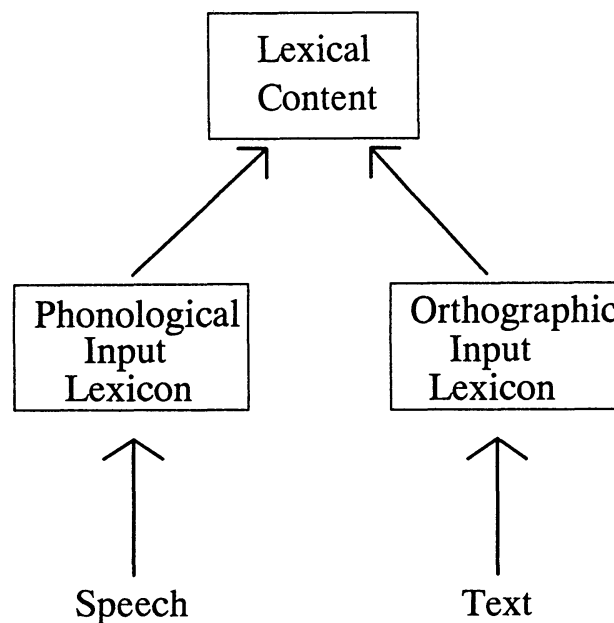

Figure 1: A generic model of lexical organisation, with functionally and architecturally parallel input from speech and text.

There are two principal claims being made in this picture, and in pictures like it. The first is *architectural*: that the basic organisation of the language system distinguishes modality-independent representations (the common core of word meaning) from modality-specific but architecturally parallel access systems. The second is *functional*: that there are basic functional parallelisms between the two modalities, so that lexical representations are accessed by qualitatively similar procedures, irrespective of the modalities involved.

We will try to establish in this paper that both these claims are wrong, and that access from the auditory, or phonological, route is neither architecturally nor functionally equivalent to access from the visual, or orthographic, route. There are three components to our claims here:

(i)   That central representations are not, strictly speaking, modality independent. Rather, we are dealing here with a level of representation we can call the *lexical entry*, where this is a triplet of abstract semantic, syntactic, and phonological information.

(ii) That phonological inputs can map directly onto these underlying, central representations, without the need for intermediate access representations. In fact, as we will argue, this involves *direct access* to decomposed morphemically organised representations.

(iii) That orthographic inputs do not have the same kind of direct access to the lexical entry, and access may well involve some form of *mediating* representation.

We will organise the argument for this view around the evidence for these last two claims, beginning with the evidence for direct auditory lexical access, and then going on to examine parallel sets of experiments in the visual domain, revealing an input system with quite different properties. In all of these experiments we rely on morphologically complex words. To make explicit the internal properties of the English mental lexicon, as an abstract, combinatorial, and morphemically organised system, it is necessary to use stimulus materials that allow us to dissociate morphemes from words, and surface form from underlying representation.

2. DIRECT ACCESS FROM SPEECH

The first part of our argument is that phonological inputs can map directly onto morphemically organised representations at the level of the lexical entry, where these representations include an abstract specification of the phonological properties of the words involved, and where the term 'phonological input' refers to the acoustic-phonetic analysis of the speech input stream, most likely delivered to the lexical level in featural form (for a closer look at these issues, see Marslen-Wilson & Warren 1994). The evidence for direct mapping comes, in the first instance, from experiments involving morphologically related words where the relationship between these words varies in its phonological transparency. Table 1 lists a set of experimental contrasts along these lines, designed for use in repetition priming experiments.

In our original experiments (Marslen-Wilson, Tyler, Waksler and Older 1994) we used a cross-modal version of the priming task, where a spoken prime word - for example, *friendly* - is immediately followed by a visually presented target word - for example, *friend* - to which the listener has to make a speeded lexical decision response (that is, decide as quickly as possible whether the target item is a word or not). We find in this task that response times are generally faster when the target