Abstractness and transparency in the mental lexicon

William Marslen-Wilson, Lorraine Komisarjevsky Tyler, Rachelle Waksler, & Lianne Older

Birkbeck College, University of London,
Department of Psychology,
Malet St.,
London WC1E 7HX
Email: ubjta38@cu.bbk.ac.uk

Abstract

This research is concerned with the structure and properties of the mental representations for morphologically complex words in English. In a series of experiments, using a cross-modal priming task, we ask whether the lexical entry for derivationally suffixed and prefixed words is morphologically structured or not, and how this relates to the semantic and the phonological transparency of the relationship between the stem and the affix (*govern* + *ment* is semantically transparent, *depart* + *ment* is not; *happy* + *ness* is phonologically transparent, *vain* + *ity* is not). We find strong evidence for morphological decomposition, at the level of the lexical entry, for semantically transparent prefixed and suffixed forms, independent of the degree of surface transparency in the phonological relationship between the stem and the affix. Semantically opaque forms, in contrast, seem to behave like monomorphemic words. We discuss the implications of this for a theory of lexical representation and the processes of acquisition.

Introduction

To understand language comprehension we need to understand the mental lexicon. This requires us to understand how words are *represented* in the mental lexicon. What are the processing targets of lexical access, and how are representations of lexical form and of lexical content interleaved to permit access, selection, and integration to operate in the rapid and efficient manner documented in earlier research (Marslen-Wilson, 1989; Marslen-Wilson & Tyler, 1987)?

The research we report here is concerned with the properties of the underlying *unit* of lexical representation. Are words represented in the mental lexicon as complete, unanalysed word-forms -- corresponding, more or less, to the words we see on the printed page? Or is the representation broken down into *morphemes* -- traditionally, in linguistic analysis, the smallest meaning-bearing linguistic unit. The word *happy*, for example, is a single morpheme, whereas the words *happiness* and *unhappy* are polymorphemic, being made up, respectively, of the morphemes {happy} + {ness} and {un} + {happy} (where {happy} is the *stem*, and {-ness} and {un-} are morphological *affixes*). In a series of experiments, using the cross-modal repetition priming technique, we explore two basic questions about the representation of morphologically complex words in English: Is their representation phonologically abstract, and what is the role of semantic factors in determining whether or not a word will be represented as morphologically complex?

Preliminaries

In studies of lexical representation, it is crucial to distinguish claims about the *lexical entry* for a given word from claims about its *access representation*. The lexical entry we define as the modality-independent core representation of a word's abstract syntactic, semantic, and phonological properties. The access representation we define as the modality-specific perceptual target for lexical access, constituting the route whereby information in the sensory input is linked to a given lexical entry. Our concern here is with the properties of the lexical entry,

and with the role of *semantic* and *phonological transparency* in determining its properties.

These terms refer to two aspects of the surface relationship between stems and affixes in morphological complex words. Semantic transparency refers to whether or not the meaning of a morphologically complex word is synchronically derivable from the meaning of its parts. Words like *happiness* and *unhappy*, for example, are relatively semantically transparent, because their meaning is composable in this way. In contrast, words like *department* or *release* are not semantically transparent -- the meaning of *department* cannot be derived by putting together the meaning of the free stem *depart* with the affix *-ment*. Our research will investigate the role that this factor, of surface semantic interpretability, plays in determining whether or not the lexical entry for a given word-form is morphologically structured.

The second factor, of phonological transparency, refers to the degree to which processes of phonological alternation lead to a change in the phonetic realisation of the stem when it occurs in a morphologically complex word. In English, this applies in particular to the suffixing morphology, as in alternations like *vain/vanity* or *decide/decision*. If pairs like *decide/decision* do share the same stem at the level of the lexical entry -- i.e., the morpheme {decide} -- then this must be represented in a way which abstracts away from the surface phonetic properties of the word-forms in question.

To investigate these issues we will focus on English *derivational* morphology, since this provides the appropriate range of phonological and semantic contrasts. Derivational morphemes in English are both suffixing and prefixing. Suffixing morphemes like *-ness, -ment, -ence* come after the stem (as in *happiness, government, defiance*, etc), whereas prefixing morphemes like *re-* , *ex-*, *pre-* (as in *rebuild, explain, preview*, etc) precede the stem. These are all morphemes which function to change the meaning, and often the grammatical class of the stems to which they are attached.

The experimental task we will use is *cross-modal immediate repetition priming*. This is a task in which the subject hears a spoken prime -- for example, *happiness* -- and immediately at the offset of this word sees a visual probe -- for example, *happy* -- which is related in some way to the prime. The subject makes a lexical decision response to this probe (i.e., judges as quickly as possible whether the string of words presented constitutes a word or not in the language). Response latency relative to a control condition, where listeners respond to the same probe following an unrelated prime, is used to measure any priming effect. Because the task is cross-modal, any priming effects should be attributable to events at the level of the lexical entry, rather than to effects of lower-level overlap at the level of modality-specific access representations.

Abstractness in lexical representation

A theory of lexical representation which claims that morphologically related words share the same stem morpheme in the lexical entry will need to assume that this is a level of representation which abstracts away from surface variation in phonological form. We test this in Experiment 1 by varying the phonological and morphological relationship between the auditory prime and the visual probe.

Note that morphological relatedness is defined here strictly on linguistic and historical grounds. A derived form and a free stem are classified as morphologically related if (a) the derived form has a recognisable affix; (b) when the affix is removed the resulting stem is the same as the paired free stem; (c) the pair of words share the same historical source word (or *etymon*).

In Condition 1 (see Table 1), the auditory prime (always a derived suffixed word) is morphologically related to the visual target (always a free stem) and this relationship is phonologically transparent. Examples of this are pairs like *friendly/friend* or *government/govern*, where the stem is fully contained within the derived word, in a form which is phonologically identical to its realisation as a free form.

If the lexical entries for words like this are morphologically decomposed into stems and affixes, then priming should be obtained in Condition 1 on the basis of shared morphemes in the lexical entry. Hearing *government* should activate the stem morpheme {govern} and the link between this and the suffix {-ment}. If the same stem functions as the lexical entry for the morphologically simple form *govern*, then residual activation of this morpheme after *government* has been heard should facilitate lexical decision responses when *govern* is presented as a visual probe.

Conditions 2 and 3 present the subjects with prime/target pairs which are still morphologically

Table 1: Phonological Transparency and Morphological Relatedness

	Test	*Control*	*Difference*
Condition 1 (friendly/friend)	539	583	-44
Condition 2 (elusive/elude)	563	623	-60
Condition 3 (vanity/vain)	572	608	-36
Condition 4 (termite/term)	647	638	9

related, but where this relationship is no longer phonologically transparent. In Condition 2, we used cases like *tension/tense* or *elusive/elude*, where the phonetic form of the stem is different in isolation from what it is in the derived form. If priming is due to events at the level of the lexical entry, then changes in the surface relationship between forms should not reduce the amount of priming. Condition 3 uses pairs like *vanity/vain* or *gradual/grade*, where not only does the stem have a different phonetic form in isolation, but also the underlying representation of the stem is not identical to its surface form. This has the effect of increasing the abstractness of the relationship between the stem and the phonetic form of the derived word.

The lexical decision responses, given in Table 1, show significant amounts of priming in the three conditions with morphologically related primes and targets. For each of these conditions, responses are significantly faster following the test prime than the control prime, and the size of the facilitation effect does not differ statistically across the three conditions.

There is no sign here that the effectiveness of a prime depends on the surface phonological transparency of the relationship between prime and target. To the contrary, when there is *only* a phonological relationship between prime and target, as in Condition 4, then no priming is obtained. Pairs such as *termite/term* or *planet/plan* are not morphologically related, so that there is no shared morpheme in common. Thus, although the target is transparently contained within the prime, there is no priming at the level of the lexical entry -- which is evidently the level of the system being tapped into by the experimental paradigm.

These results not only support the view that lexical representations are morphologically structured, but also that these representations are abstract. At the level of the lexical entry, representations of lexical form do not simply reflect surface form. If they did, then *decision* would be a much less effective prime of *decide* than *friendly* would be of *friend*, where the surface phonetic overlap is much greater.

Semantic Transparency

What is the role of semantic factors in the priming that we observed in Experiment 1? All of the morphologically related pairs (Conditions 1-3) were semantically transparent, whereas pairs like *planet/plan* (Condition 4) clearly were not. It is possible, therefore, that the lexical relations we are tapping into are semantic in nature and not necessarily morphological at all. The words *government* and *govern*, for example, share many semantic properties, and it may be by virtue of this relationship, rather than any specifically morphological relationship, that priming is obtained. This raises the issue of whether there are grounds for supposing that there is morphological structure in the lexicon independently of semantic structure.

The clearest arguments here are linguistic in nature. Aronoff (1975), for example, argues that morphological relations can be identified which involve morphemes that have no clear semantic interpretation. These are cases like the bound morpheme {-mit}, which only occurs as an element in words like *permit*, *transmit*, and *submit*. Although these words do not share a common meaning, they are linked by a common phonological rule, which generates the forms *permission*, *transmission*, and *submission*, and which is specific to verbs containing the root {-mit}. This suggests, according to Aronoff, that phonetic strings can be identified as morphemes independently of semantic considerations.

Returning to English derivational suffixes, there are plenty of cases where morphological links can be established between pairs of words, but where the relationship is no longer semantically transparent. These are cases like *emergency/ emerge* or *department/depart*, which meet the criteria for morphological relatedness, but where the meaning of the complex form can no longer be derived from the simple composition of the meanings of the stem and the affix. In Experiments 2 and 3 we contrast priming for semantically unrelated but morphologically related

Table 2: Semantic Transparency and Morphological Relatedness

		Test	Control	Difference
Suffixes:	Semantically Related (punishment/punish)	554	595	-41
	Semantically Unrelated (department/depart)	575	574	1
Prefixes:	Semantically Related (insincere/sincere)	503	534	-31
	Semantically Unrelated (restrain/strain)	542	543	-1

pairs with priming for semantically and morphologically related pairs of the type used in Experiment 1, such as *friendly/friend* or *predictable/predict*. Experiment 2 looked at suffixed forms and Experiment 3 at prefixed forms (Table 2).

These contrasts require an assessment of semantic transparency for each test pair. Derived forms are transparent when the meaning of the form is largely determinable from the composition of the stem (or root) with its affix. To determine whether this held synchronically -- that is, for current users of the language -- we used an operational criterion, classifying words as semantically transparent or opaque on the basis of a pre-test, where individuals were asked to judge the relatedness of a derived form and its free stem.

The lexical decision responses (Table 2) show clear effects of semantic relatedness. Although all pairs were morphologically related, according to uniformly applied linguistic and etymological criteria, only those pairs that were also synchronically semantically related showed priming in this task. In other experiments (Marslen-Wilson, Tyler, Waksler, & Older, 1992) we have found the same pattern when the order of prime and target is reversed -- *punish* is a strong prime of *punishment*, but *depart* does not prime *department* (and similarly for prefixed pairs). We also find no evidence of priming for prefixed pairs that share bound stems (of the *include/conclude* type), where there is again morphological but no synchronic semantic relation.

In a final set of experiments (Table 3) we examine the priming effects for suffixed and prefixed pairs where both prime and target are derived forms, sharing the same stem. If priming in this task is due to residual activation of a shared morpheme in the prime and the target, then semantically transparent pairs like *government/governor* should prime each other, for the same reason that *govern* primes *government*, and *vice versa*.

What we find (Table 3) is that although prefixed derived pairs do prime each other, suffixed pairs do not. Semantic relatedness is therefore not enough on its own to ensure priming between morphologically related pairs. We can attribute this absence of priming to inhibitory links between suffixes sharing the same stem. Hearing a transparent suffixed form like *government* will not only activate the stem {govern} but also inhibit other suffixed forms sharing the same stem. This is because forms like *government* and *governor* are mutually exclusive competitors for the same lexical region. The combination of the morpheme {govern} with the affix {-ment} defines a lexical item with a distinct meaning and identity in the language, and this is incompatible with the simultaneous combination of {govern} with a different affix to give a different lexical item.

The fact that prefixed derived pairs (such as *unwind/rewind*) do prime each other (Table 3) is evidence that the lack of priming for suffixed pairs is indeed a competition effect. When the lexical entry is entered *via* a prefix, this does not seem to activate as competitors other prefixed words sharing the same stem. In effect, *rewind* is not in the same cohort as *unwind*, so that these competitors do not need to be inhibited in the same way as suffixed words sharing the same stem.

Table 3: Semantic and Morphological Relatedness: Derived/Derived Pairs.

		Test	Control	Difference
Suffixes	Semantically Related (excitable/excitement)	580	591	-11
	Semantically Unrelated (successful/successor)	611	614	-4
Prefixes	Semantically Related (unfasten/refasten)	576	635	-59
	Semantically Unrelated (express/depress)	576	554	22

Conclusions

This series of experiments allow us to draw three main conclusions:

(i) There is a level of lexical representation which is abstract in nature. Phonetic overlap between primes and targets does not by itself produce priming, and the amount of priming is not affected, for morphologically related forms, by variations in the phonological transparency of the relation between prime and target.

(ii) Semantic relatedness between a prime and a target is a necessary but not sufficient condition for priming to occur. Semantically unrelated pairs, whether morphologically related or not, do not prime reliably.

(iii) The type of morphological relation between a prime and a target, and the stem-affix order within a morphologically complex prime, affect whether or not priming is obtained.

To accomodate these results we need to postulate a model of the mental lexicon which treats separately words like *department* (which are semantically opaque) and words like *punishment* (which are semantically transparent) at the level of the lexical entry. Semantically opaque words will be represented as if they were morphologically simple -- they can enter into combination with other morphemes (as in *interdepartmental*) but they themselves have no internal structure. Synchronically transparent forms, in contrast, will be represented as free stems linked to derivational affixes. Within this system of linked stems and affixes, inhibitory links will need to be set up between suffixes sharing the same stem, but not between prefixes, or, indeed, between prefixes and suffixes.

This proposal has the effect of re-interpreting semantic relatedness in terms of its consequences for the learning process. The structure of the adult lexicon reflects individuals' experience with the language as they learn it. The listener does not mentally represent words as sharing the same stem, and therefore as morphologically related, unless there are semantic grounds for doing so. An item like *department*, although it has a phonetically transparent morphological structure on the surface, will not be analysed during language acquisition into the free stem {depart} plus the affix {-ment} at the level of the lexical entry, since this gives the wrong semantics.

The challenge for learning models of English derivational morphology will be to devise ways of allowing phonological and semantic criteria to interact in the complex but rational ways that this research has begun to uncover.

Acknowledgements

This research was supported by grants to William Marslen-Wilson and Lorraine Komisarjevsky Tyler from the UK SERC, ESRC, and MRC. Rachelle Waksler at now in the Department of Linguistics, San Francisco State University, San Francisco.

References

Aronoff, M. (1976) *Word Formation in Generative Grammar*. Cambridge, Mass.:MIT Press.

Marslen-Wilson, W.D. (1989) Access and integration: Projecting sound onto meaning. In W.D.Marslen-Wilson (Ed.), *Lexical Representation and Process*. Cambridge, Mass.: MIT Press.

Marslen-Wilson, W.D., & Tyler, L.K. (1987) Against modularity. In J.Garfield (Ed.), *Modularity in Knowledge Representation and Natural-Language Understanding*. Cambridge, Mass.:MIT Press

Marslen-Wilson, W.D., Tyler, L.K., Waksler, R., & Older, L. (1992) Morphology and meaning in the English mental lexicon. Manuscript, Birkbeck College.