Chapter 7

Modelling the Processing of Morphologically Complex Words

R. Harald Baayen and Robert Schreuder

Introduction

This chapter reviews a number of linguistic and psycholinguistic computational models of morphological processing. Morphologists study the internal structure of words. For instance, a word such as *antidisestablishmentarianism* can be assigned the following structure:

$$[_N\text{anti} + [_N\text{dis} + [_N[_N[_A[_N[_V\text{establish}_V] + \text{ment}_N] + \text{ary}_N] + \text{an}_N] + \text{ism}_N]_N]_N]$$

For English, words with such a complex internal structure are rare. In languages such as Turkish or Greenlandic Eskimo, however, very complex words are far more common. The question with which we are concerned is how the human language processor deals with morphologically structured words of varying complexity. What kinds of mechanism allow listeners to interpret novel as well as familiar complex words? What kind of experimental evidence is available to constrain computational models? What types of morphological operations are attested in human languages? And what requirements do these operations impose on computational models? In this chapter, we first survey the basic linguistic facts that any model should attempt to account for. The following section describes two linguistic computational models, Koskenniemi's (1983) two-level morphology model and Ritchie *et al.*'s (1992) unification-based model. After that, the results of a number of psycholinguistic experiments are summarized, followed by an outline of a number of verbal (not implemented) models of morphological processing, as well as Gasser's (1994a, 1994b) connectionist model for the processing of complex words in the auditory modality. We will see that the complexity of the phenomena involved is such that none of the existing psychological models are able to handle all aspects of morphological processing, and that they are even less

able to predict actual experimental reaction times. These and other problems are discussed in the last section.

Some Basic Linguistic Facts

Languages exploit a variety of means for creating morphologically complex words. Often, the means by which morphologically complex words are created involve more complex operations than the simple linear concatenation of morphemes. Hence the modelling of the mapping of form onto meaning in perception cannot depart from the assumption of a simple one-to-one relation between input representations and central representations (*templates*) in the mental lexicon. We review the main linguistic devices by which semantic or syntactic information is morphologically expressed, and discuss the role of phonology. Some of the complexity issues that may be encountered are briefly summarized. Finally, we focus on the semantic aspects of word formation. Throughout our discussion we trace the processing consequences of the linguistic facts discussed.

Morphological Operations

Broadly speaking, we may distinguish between two fundamentally different morphological processes: concatenative and non-concatenative operations. Concatenative operations involve the stringing together of free and bound morphemes. Generally, we distinguish between prefixation, e.g., *re-forest*; suffixation, e.g., *forest-ation*; circumfixation, e.g., Dutch *ge-wandel-d* (past participle of *wandelen*, 'to walk') and compounding, e.g., *door-bell*. The complexity of the resulting words may vary considerably. In English the total number of morphemes that are likely to appear together in one complex word is rather limited. The example in the introduction, from Sproat (1992, p. 44), concerns an English word with seven morphemes, but the maximum number of morphemes registered in the CELEX lexical database of English equals only five, e.g., *de-nation-al-ize-ation*. By way of contrast, Turkish (Altaic) allows words such as *avrupa-lI-las-tIr-Il-a-mI-yacak-lar-dan-sIn-Iz-dIr*, 'surely you (all) are among those who will not be able to be caused to become like Europeans', with 13 morphemes (Beard, 1992, p. 96) and, somewhat less extreme, *göz-lük-çü-lül-çük-çü*, 'proponent of optometry proponency' (Beard, 1992, p. 94), with six morphemes. Note the affix-based recursion in the last example (. . . *çük-çü* . . .), something that leads to a combinatorial explosion of possibilities that is rarely encountered in Indo-European languages. On the other hand, compounds in English may be of a similar degree of complexity, e.g., *computer communications network performance analysis primer* (Sproat, 1992, p. 41).

Non-concatenative morphological operations involve infixation, root-and-pattern morphology, subtractive morphology, reduplication, subsegmental and suprasegmental morphology and zero-marking. In Tagalog (Austronesian), the infix *-in-* denotes the object which is the result of some action. Characteristically, this infix is inserted not after a morpheme boundary but right after the first consonant of the stem: *sulat*, 'a writing', *s-in-ulat*, 'that which was written' (Schachter and Otanes, 1972).

Root-and-pattern morphology is typical for the semitic languages. It mostly occurs in conjunction with concatenative processes. Consider the Hebrew verb root *šmr*, 'to guard'. In Classical Hebrew, this purely consonantal root is fused with the vowel pattern *a-a* in *šamar-ti*, 'I have guarded', but with a single stem vowel (*-o-*) in *e-šmor*, 'I will guard' (Gesenius and Kautzsch, 1909). These changes in the consonantal and vowel patterns require a fundamentally different parsing technique, even though this kind of morphology can be analyzed as involving concatenation of morphemes in a multi-dimensional representational system (McCarthy 1982).

Subtractive morphology involves the deletion of part of the base in order to mark some semantic operation. In Koasati (Algonquian, Northern Amerindian), to pluralize the subject of a verb one has to delete the last rhyme of the stem: *obakhitip-li-n*, 'he goes backwards', *obakhit-li-n*, 'they go backwards' (Martin, 1988). Note that the rhyme, and hence the deleted segment, will vary from word to word. The absence of a constant segment to meaning association poses a serious problem for any computational theory of word processing (Sproat, 1992).

Reduplication involves the partial or total repetition of a morpheme. Malay (Austronesian) uses total reduplication to express the plural: *kursi*, 'chair', *kursikursi*, 'chairs'. Partial reduplication processes typically depend on the phonological structure of the base word. Samoan (Austronesian) reduplicates the next to last syllable to express plural subjects on verbs: *savali*, 'he travels', *savavali*, 'they travel' (Gleason, 1955).

Subsegmental morphological processes involve the expression of some semantic function by changing some features of a segment of the base. In Nuer (Nilo-Saharan), the phonological feature [*continuant*] distinguishes the negative form of the past participle from its positive counterpart: *jaac* is the negative form of the present participle of the verb 'to hit', of which the past participle is *jaaç* (Lieber 1987).

As an example of a semantic function that is marked by suprasegmental means consider Mono-Bili (Niger-Kongo). In this language the future tense is expressed by a low tone on the verb. The past tense is expressed by a high tone. Thus we have *mṹrṹ wó sè*, 'the leopard kiiled him', which contrasts with *mṹrṹ wò sè̃*, 'the leopard will kill him' (Jensen, 1990, p. 74).

Finally, semantic changes are often accompanied by no morphological marking at all. Lieber (1992) argues that two cases of zero-morphology should be distinguished. On the one hand, French instrument-agent nouns such as *essuie glace*, 'windshield wiper', may well involve a semantically specified but phonologically unrealized suffix, given the uniformity of the interpretation of

these compounds and the gender they select (masculine: *le*). On the other hand, noun to verb conversion in English, *to book*, expresses such a wide range of meanings that it is implausible to assume that a single (zero) affix is involved. (But see Beard, 1992, for a different view.)

Phonological Issues

When morphemes are combined, various phonological processes may alter their segmental and suprasegmental make-up. Since such processes may significantly complicate the perception process, we present a brief overview of the kinds of change one may encounter.

A first set of changes concerns the form in which particular segments appear. Often the segments at the boundaries of morphemes undergo assimilation (*in-dubitable, im-proper, in-glorious*, the latter pronounced /ŋ/). But assimilation processes may affect segments at longer distances too. A well-known example is vowel harmony. In Hungarian (Finno-Ugric), the exact quality of the vowel of the suffix *-nek/-nak* is determined by the vowel quality (front or back) of the stem. Thus we have *öröm-nek*, 'to joy', but *hấz-nak*, 'to house' (Jensen, 1990, p. 163). Furthermore, consonants and vowels may be inserted (epenthesis). For instance, in order to avoid a sequence of consonants that is difficult to pronounce, the Latin stem *patr-*, 'father', as in *patris*, 'father (genitive case)' appears with an extra vowel (*pater*) in the nominative case. Segments may also interchange (metathesis), as for instance in Hanunoo (Austronesian), where multiplicative adverbials like 'once' and 'twice' are derived from the corresponding numerals by prefixing *ka-* and interchanging a stem-initial glottal stop with the following consonant: *ka + 'sa → kas'a*, 'once', where *-'sa* is the root of 'one' (Schane, 1973, p. 56). Another kind of phonological change that should be mentioned is final devoicing in some Germanic languages, where stem-final voiced stops and fricatives become voiceless if no other affixes follow: compare Dutch *huizen*, 'houses' with *huis*, 'house'.

A second set of changes concerns the form in which an affix or stem appears. The English nominalizing suffix *-ion* appears as *-ation*, *-ution*, *-ition*, *-tion* and *-ion*, depending on the phonological properties of the stem, although the choice is often morpholexically conditioned (for discussion see Aronoff, 1976). In addition to allomorphy of the affix, allomorphy of stems may complicate matters further. For instance, the Latin stem *cap-* as in *capio*, 'I seize', appears as *cep-* in the perfect: *cepi*, 'I have seized' (for a detailed analysis of stem allomorphy see Aronoff, 1994).

Some of these phonological changes may be perceptually or articulatorily motivated (Schane, 1973, p. 61). For instance, the insertion of a vowel in the nominative of Latin *patr*, 'father', may be articulatory functional, but it complicates the mapping of the various forms in which such stems appear onto their corresponding lexical representations. Similarly, the mismatch between syllable structure and morphological structure for words like Dutch *masten* (/mɑs$tən/ — *mast-en*), 'masts' might be dysfunctional for speech perception, where the perceptually most salient units /mɑs/ and /tən/ do not map onto the morphemes *mast* and *-en*.

Complexity Issues

Let us assume that the perceptual system solves the various phonological and non-concatenative complexities discussed above, and maps the speech input onto the constituent morphemes of a given complex word. At this point problems of another nature may be encountered: how to process long-distance dependencies, how to decide between alternative parsings, and how to handle affixal homonymy.

At first sight, it would seem possible to process a string of morphemes one by one, rejecting a string as soon as two morphemes cannot be legitimately combined. Such a processing algorithm fails, however, for words with so-called long-distance dependencies. Consider the Dutch adjective *on-denk-baar*, 'unthinkable'. Crucially, the prefix on- does not attach to verbs: the hypothetical verb *on-denk-en*, 'to unthink', is ungrammatical. It is only after the verb stem *denk* has been integrated with the affix *-baar*, '-able', to result in the adjective *denkbaar* that *on-* is licensed for combination. This example shows that simple left-to-right processing without some form of buffering and look-ahead is doomed to fail.

A problem specific to the recognition of complex words is that often large numbers of possible and competing structural analyses can be assigned. For instance, the Dutch word *belangstellende*, 'interested person', should normally be parsed as *belang-stel-en-d-e*, 'interest-put' followed by suffixes for infinitive, participle, and adjectival inflection, but the reading *bel-angst-ellende*, 'misery of having fear for ringing up' is also possible. According to Heemskerk (1993, personal communication), orthographic complex words in Dutch have on average some three different parsings. The number of different parsings is somewhat reduced for the auditory modality, where stress and syllabification patterns and vowel quality may rule out certain parsings, as for the example given above (/*bə$lɑŋ$'stɛl$lən$də*/ or /'*bɛl$ɑŋst$ɛl$lɛn$də*/). It may also happen that a particular string may be either a monomorphemic word or a complex word. For instance, Dutch monomorphematic *beton* generally denotes the building material 'concrete', but phonologically and orthographically identical *be-ton*, 'to mark with buoys', is a less likely but nevertheless established nautical term. Frequency of use and contextual information are crucial for solving ambiguities of this kind.

Affixal homonymy poses yet another problem. In Dutch, the suffix *-te* may attach to verbs to express the singular past tense, e.g., *zwalk-te*, 'roamed', or it may attach to adjectives to form abstract nouns, e.g., *zwakte*, 'weakness'.

According to Beard (1992, p. 78), it is quite common to find productive affixes that serve both derivational and inflectional duties. Within inflection, one and the same affix may be polyfunctional too (*-s* in Dutch and English expresses either the plural or the genitive), and the same holds for derivation (*in-* in Dutch and English expresses either direction, e.g., *intake*, or negation, e.g., *insecure*). Thus identifying the right form does not necessarily lead to a unique correct meaning. The correct meaning can only be established on the basis of the subcategorization properties of the affixes and stems involved.

Semantic Issues

Many complex words, however, are in some way irregular. Such irregularity is especially pervasive at the semantic level in derivation and conversion. One may observe degrees of semantic transparency, ranging from full transparency as in *goodness* to complete opacity, as in *disease*. However, even the slightest unpredictable property of the meaning of the complex word requires the storage of such a property. The number of words with a particular affix that require such storage may be quite substantial. Schreuder and Baayen (1994) observed for a selection of Dutch prefixes that the number of types with at least partially unpredictable readings may amount to 48 per cent of all types with that prefix. In terms of frequencies of use, roughly 70 per cent of all corresponding tokens turn out to be opaque. Such corpus-based calculations show that it may be disadvantageous to force a parsing on each and every input.

Another important distinction is between semantic transparency and semantic complexity. Among the transparent formations, some involve more complex semantic operations than others. For instance, inflectional processes such as pluralization involve a relatively simple semantic operation on the meaning of the stem. Deverbal verb forming processes, on the other hand, may require fairly complex operations on the lexical conceptual structures (Jackendoff, 1991; Lieber and Baayen, 1994) of base and affix. For instance, the Dutch verb *bouwen*, 'to build', is subcategorized for a direct object, the thing being built. Prefixation with *be-*, as in *bebouwen*, 'to build up', gives rise to a verb in which the location on which something is being built appears as the direct object. Such changes in valency patterns introduce additional computational complexity.

Summary

Summing up, the basic linguistic constraints on any computational model of lexical processing are the following. First, the variety of means available for the creation of complex words, both concatenative and non-concatenative, shows that a simple template matching view of morphological processing is inadequate. In as far as these morphological processes are productive, i.e., that they lead to the creation of new words that can be readily understood, they should be taken seriously by any computational model. Second, morphological operations on words may be sensitive to phonological structure. A computational model of lexical processing must include representations for syllables, rhymes, onsets, etc. in addition to bare phoneme sequences. We shall see that this requirement poses considerable problems for almost all computational models. Third, the mapping of form representations on the associated meaning representations may be obscured by phonological processes such as epenthesis and assimilation. Computational models should be able to handle such phonologically conditioned alternations. Fourth, preferential parsings should be given priority in computational models, without giving up the availability in principle of competing, less likely segmentations. Finally, the extent to which semantic opacity pervades the lexicon suggests that lexical storage often takes precedence over parsing.

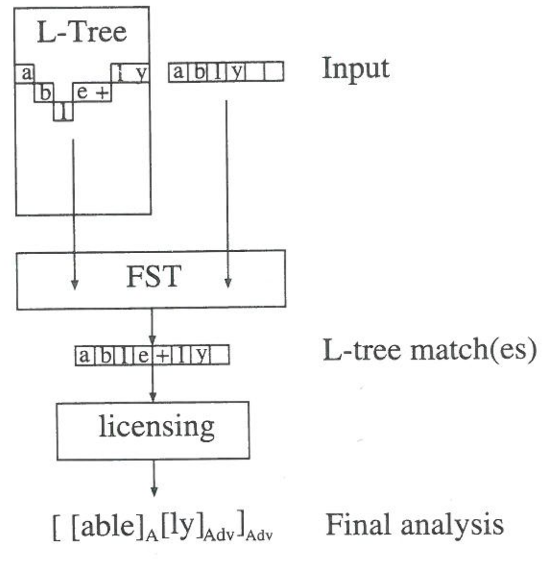
Linguistic Models

We now turn to describe two operational computational parsing models developed by computational linguists. It is instructive to consider these models in some detail, as they implement efficient algorithms that actually perform the task for which they are designed. By studying these algorithms, one's insight into the complexities involved is enhanced. At the same time, various aspects of these models are relevant for computational modelling in psycholinguistics. In what follows, we describe morphological processing as consisting of three interrelated stages: segmentation (discovering the morphemes), licensing (checking for compatibility of subcategorization), and composition (computation of the syntactic and semantic structure of the complex whole).

Most current linguistic computational models of morphological parsing are designed as shown in Figure 7.1.

Parsing is accomplished in two steps. First, in the segmentation stage, the input is compared with underlying lexical representations in the lexicon. The system attempts to match the input to one or more lexical representations, taking into account phonologically or orthographically determined differences between surface and underlying forms. For instance, the orthographic input *ably* is recognized as *able + ly*. The result of this first step is a set containing one or more flat segmentations, that is, strings of morphemes which match the input given the orthographical and phonological rules of the language. No hierarchical structure is as yet assigned. As we shall see below, the lexicon is assigned a tree-like structure (the so-called L-tree, trie or sieve), and the segmentation process is carried out by a

Figure 7.1: General architecture of linguistic computational models



particular kind of finite-state mechanism, the so-called finite state transducer (FST). The first operational parser of this type was called KIMMO, after its creator, Kimmo Koskenniemi (1983). This approach to segmentation is generally known as *two-level morphology*, but might just as well be described as *two-level phonology* (cf., Anderson, 1992).

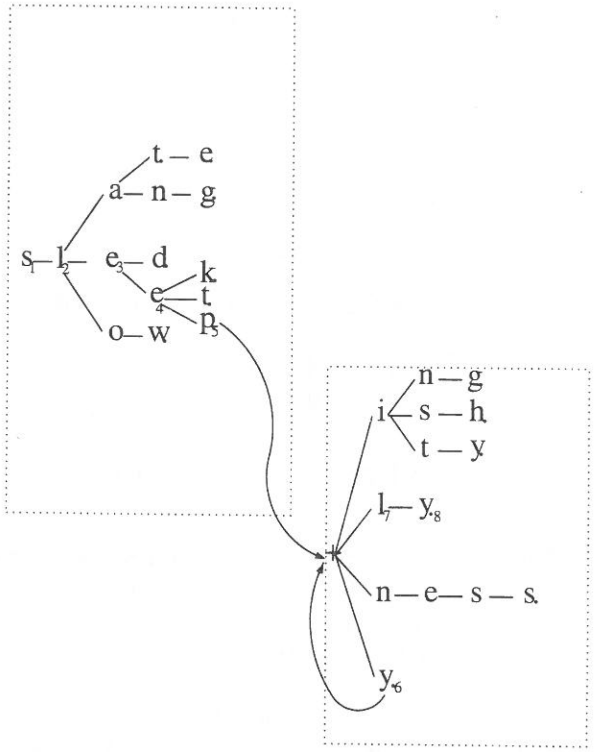
Once the set of possible flat segmentations has been obtained, the next step, licensing, is to select those segmentations that do not violate the subcategorization properties of the morphemes involved. For instance, the Dutch string *beneveling* can be assigned six flat segmentations, each consisting of a series of legitimate morphemes: *be + neef + eling*, *be + neef + eel + ling*, *be + neef + e + ling*, *been + e + veel + ing*, *be + en + e + veel + ing*, and *be + nevel + ing* (Heemskerk, 1993). Only the last flat segmentation is licensed by the subcategorizations of all morphemes involved, leading to the meaning 'misting up'. Thus the goal of the parsing process is to obtain one (or more) legitimate labelled bracketing(s) for the input, often in combination with a specification of the subcategorization and other syntactic and semantic properties of the analyzed complex word (composition).

In the next two sections we discuss these two steps in more detail.

Segmentation

In KIMMO-like parsers, segmentation is accomplished by means of an *L-tree* or *trie*

Figure 7.2: Scanning the input string sleepily *using an L-tree. The numbers show the successive steps through the L-tree*



(Knuth, 1973) and a *two-tape finite state transducer* (Koskenniemi, 1983, 1984). An L-tree is a tree structure in which each node is labelled with a particular grapheme or phoneme. To ascertain whether a given string is stored in the L-tree, the input string is scanned from left to right, each successive character being matched with a character in the L-tree, starting at the root. This scanning process is shown for the input string *sleepily* in Figure 7.2. Beginning with the initial *s* matching the letter of the root node of the (partial) L-tree shown here, the successive letters scanned are matched with the available nodes to the right that become available as a path is traced through the L-tree. Nodes in the L-tree are marked when they specify possible word-final segments. In Figure 7.2, a dot marks these end nodes. Thus *sleep* is a legitimate word in the L-tree. In addition, end nodes may contain pointers to one or more so-called continuation lexicons. Figure 7.2 shows part of a possible continuation lexicon. This lexicon is needed for the segmentation of *sleepily*. The numbers in Figure 7.2 show the successive steps through the two L-trees that are required, leading to the segmentation *sleep* + *y* + *ly*. Notice, however, that this segmentation does not match the surface

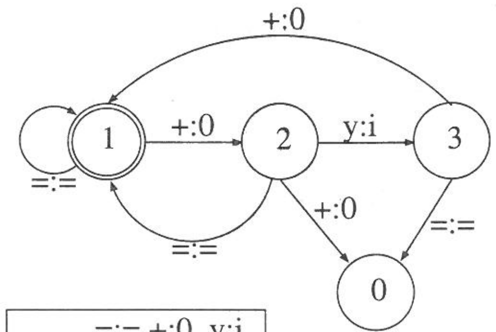
form, which does not contain any + symbols nor the grapheme *y* at the sixth position. It is here that the finite state transducers (FSTs) play a crucial role.

Two-tape FSTs are a class of automata (see Chapter 2) that operate on two sequences of symbols: the symbols of the lexical segmentations (*sleep* + *y* + *ly*) and the symbols of the surface string (*sleepily*). An FST accepts or rejects the combination of such strings. They license those mismatches that are brought about by the orthographical and phonological processes in the language. In the simplest case, each of these processes requires its own FST. All FSTs operate in parallel on the successive symbols of the two strings. One of the finite state transducers needed for accepting *sleep* + *y* + *ly* as the lexical form of *sleepily* is shown in Figure 7.3. This FST accepts surface *i* as the counterpart of lexical y, allowing the transition from the final *p* of *sleep* in the main L-tree of Figure 7.1 to the *y* in the L-tree of suffixes. The two-level rule implemented by the FST is the following expression:

(1) y:i⇒ + :0__ + :0.

This rule states that the correspondence between lexical *y* and surface *i* is legitimate only when morpheme boundaries precede and follow in the lexical representation. Note that the + :0 symbol pairs specify that a transition in the L-tree is not accompanied by a transition to the next symbol of the surface representation. A FST implementing this rule is shown in Figure 7.3, both in the form of a transition network and in the form of a transition table. State one is both the initial and the (only) final state. This state accepts any pair of symbols (= : =) until a morpheme boundary is encountered (+ :0), in which case the system enters

Figure 7.3: A finite state transducer for accepting sleep+y+ly *as the lexical form of* sleepily



	=:=	+:0	y:i
1	1	2	1
2	1	0	3
3	0	1	0

y:i => +:0__+:0

state two. If the system now encounters the pair y:i, it will enter state three. The only legitimate continuation is the pair + :0, another morpheme boundary, which will return control to the initial state. Any other pair of symbols will put the FST into state 0, the sink state, indicating that the string is not licensed by the FST. Returning to state two, we observe that the symbol pair + :0 will also lead to the sink state (the FST rejects empty morphemes). All other pairs of symbols are allowed at state two, with the FST returning to state one. The reader may verify that this FST will accept strings such as

```
sleep + y + ly     sleep + ish
sleep∅i∅ly         sleep∅ish
```

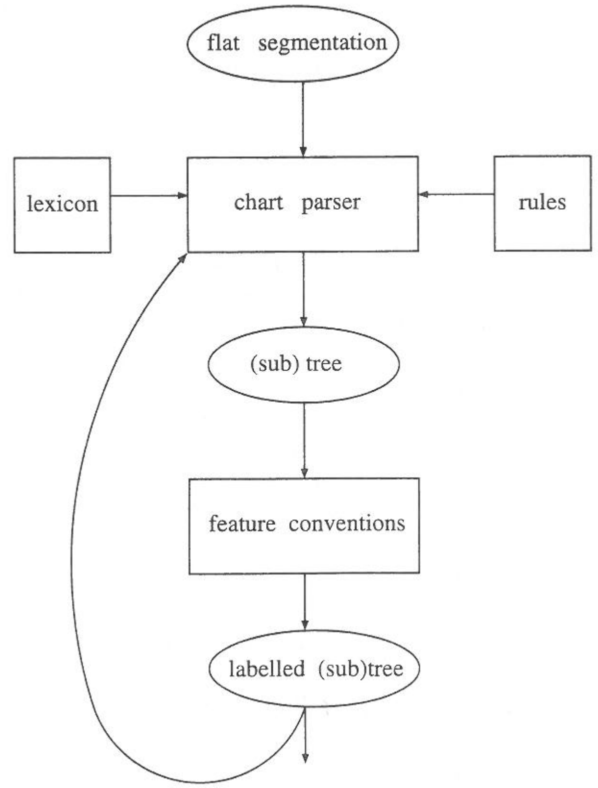
The system as shown in Figures 7.2 and 7.3 is greatly simplified. First, by using various continuation lexicons, one may model aspects of morphotactic constraints. For instance, the simplified schema of Figure 7.2 would accept the input string *sleepness*. By having different continuation lexica for the various word categories, many such strings can effectively be ruled out. Long-distance dependencies (*on-denk-baar*), however, cannot be solved in a principled way by means of linked sublexicons (see Sproat, 1992, pp. 91–2). Second, we have illustrated only how one orthographical rule operates in isolation. However, implemented two-level phonologies typically make use of 10 up to 40 FSTs operating in parallel. Special care is required for rules that interact with each other (see Sproat, 1992, pp. 132–43).

Summing up, KIMMO-based segmentation systems can be used to obtain flat segmentations of input strings into their constituent morphemes. Depending on the amount of effort put into the modelling of morphotactics by means of continuation lexicons, the number of different possible flat segmentations for a single input string may be quite substantial (see Heemskerk, 1993). Ideally, one would like a parser to select the correct segmentation from such sets of alternative parsings. In addition, the immediate constituency of the input as well as its word category and argument structure should be computed. A computational framework for accomplishing the tasks of licensing and composition has been developed by Ritchie *et al.* (1987) and Ritchie *et al.* (1992), among others.

Licensing and Composition

The general outline of the approach taken by Ritchie *et al.* (1987, 1992) is sketched in Figure 7.4. A flat segmentation is the input to a so-called *chart parser*, a parsing algorithm that systematically checks through the alternative parses, while keeping track of the previously explored subparses (see Chapter 2). The chart parser makes use of two kinds of information: a lexicon and a set of context-free rules. The lexicon is a list of the morphemes in the language. Each lexical

Figure 7.4: General outline of a system using a chart parser



entry consists of a citation field, a phonological field, a syntax field and a semantics field. The entries for the morphemes in our example of a long-distance dependency, *on + denk + baar*, 'unthinkable', would look like this:

(on, /ɔn/, [FIX PRE, BAR −1, STEM (N +)], NEGATION)
(denk, /dəŋk/, [BAR 0, N − , V +], THINK)
(baar, /ba:r/, [FIX SUF, BAR −1, N + , V + , STEM(N − , V +)], ABLE)

The relevant fields are the citation field, which contains the lexical form of the word as it may appear in the flat segmentation, and the syntactic field, which contains the so-called category of that word. Such categories are defined as unordered sets of feature-value pairs, using the so-called unification formalism. In this approach, the nodes in constituent trees are labelled by such sets of feature-value pairs and not by single terminal or non-terminal symbols (see Shieber, 1986). In the above example, the BAR feature specifies the X-bar level of a (sub)tree (cf., Lieber, 1980; Selkirk 1980). Affixes are assigned the X-bar level −1, words appear with X-bar level 0. The FIX feature specifies whether a prefix (PRE) or a suffix (SUF) is involved. Word categories are distinguished by means of the features N and V. The STEM feature specifies the word category of the stem to which an affix attaches. Ritchie *et al.* (1992) make use of a rather elaborate feature set, but this allows them to formulate a very simple word grammar. Their English word grammar contains only six context-free rules. The rules for prefixation and suffixation make crucial use of the feature-value pairs mentioned above:

[BAR 0]→[FIX PRE, BAR −1], [BAR 0]
[BAR 0]→[BAR 0], [FIX SUF, BAR −1]

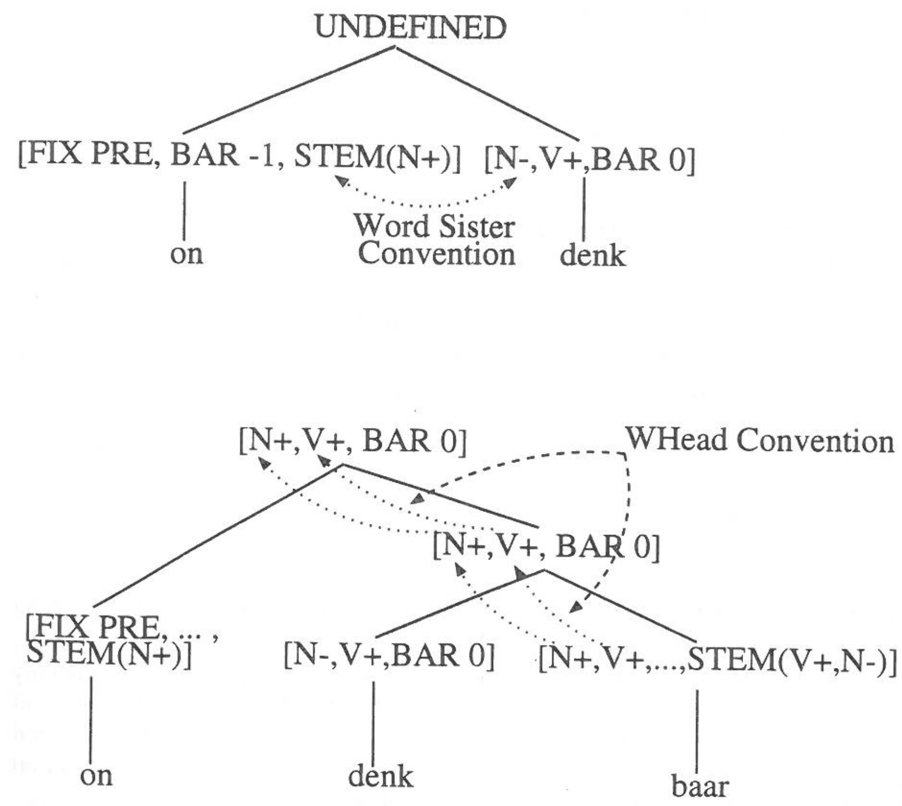
The chart parser applies such rules to its input and creates a (binary) tree structure for the first two morphemes in the flat segmentation. Subsequently, a set of so-called feature conventions is applied to this structure, enriching the newly created node with feature-value pairs and checking the compatibility of feature specifications among sister nodes. For instance, the Word-Head Convention specifies that the WHead feature-values (such as N, V) in the mother node should be the same as the WHead feature values of the right daughter. This condition is met by copying the WHead features of the right daughter to the mother. Similarly, the Word Sister Convention stipulates that when one daughter (either right or left) has the feature STEM, the other daughter should contain the feature-value pair specified as the value of the STEM feature. The application of the feature conventions results in either a legal labelled (sub)tree, or in an undefined structure. If the analysis does not exhaust the input, this result is subjected to further processing by the chart parser.

Figure 7.5 illustrates how a complex adjective such as *on + denk + baar* is analyzed. Initially, the chart parser attempts to combine the prefix *on-* with the verb stem *denk*, resulting in the uppermost tree in Figure 7.5. Subsequent application of the Word Sister Convention shows that this combination is not licensed. The chart parser then proceeds to combine the second and third segments of its input, labelling the mother node with the feature-value pair BAR 0. The Word Sister Convention is met, allowing the word category features of the right daughter to percolate to the mother node. The chart parser now combines the initial prefix with the adjective *denkbaar* [N+, V+, BAR 0], as shown in Figure 7.5. This combination is licensed by the Word Sister Convention, and the WHead features of *denkbaar* again percolate upwards by the Word Head Convention. In this way the full structure of the adjective *ondenkbaar* is obtained.

Evaluation

We have described the main features of a computational model of morphological parsing. The basic idea of a two-level approach to segmentation is widely used in current models. The unification approach to licensing and composition of Ritchie *et al.* (1992) is only one of several operational parsing techniques (see e.g.,

Figure 7.5: Word Sister Convention and Word Head Convention in Ritchie et al. *(1992)*



Heemskerk and Van Heuven's (1993) MORPA and Hankamer's (1989) KECI). However, as we see below, it is also an attractive system from a psycholinguistic point of view.

Nevertheless, a number of problems remain to be discussed. Koskenniemi's (1983) two-level morphology assumes that a direct mapping of surface forms onto lexical forms can be achieved without intervening representational levels, contrary to what most phonologists would claim. It remains to be seen whether all morphological and phonological phenomena can in fact be adequately understood without these intermediate representations (see Anderson, 1992, p. 385, but cf., Sproat, 1992, pp. 147–51).

A second set of problems concerns the number of different segmentations and alternative parsings for these segmentations with which one is often confronted. Many initial segmentations will be ruled out by the licensing conditions (feature conventions). Nevertheless, different segmentations may remain possible (*preached*, Ritchie *et al.*, 1992, p. 164). Moreover, one and the same segmentation may lead to different parse trees: the flat segmentation *be + nevel + ing* can be assigned three different labelled bracketings:

$[[be + [_N nevel]] + ing]$, $[[be + [_V nevel]] + ing]$, and $[be + [_N[nevel] + ing]]$ (Heemskerk, 1993). Probabilistic techniques for deciding between alternative analyses are now under development.

Some Basic Psycholinguistic Facts

Having reviewed the linguistic facts that constrain computational models, and having discussed some of the algorithms developed by computational linguists to meet these constraints, we now consider what additional constraints are imposed by psycholinguistic results. We discuss some main lines of evidence concerning the segmentation, licensing and composition stages. For a recent review of the experimental literature, the reader is referred to McQueen and Cutler (in press).

Segmentation

The major research thrust of psycholinguists has been aimed at establishing whether morphological structure plays a role in word recognition. In principle, morphology may be relevant to two representational levels that are generally assumed, the level of modality-specific access representations, and the level of more central, modality-free lexical representations. Most studies have addressed the issue of segmentation at the latter level. A brief overview of experimental results on morphology at this central level is presented first.

Many studies have used the repetition priming paradigm. Repetition priming involves presenting a complex word early in the experiment, to be followed by its base at a specified distance. In the baseline condition no morphologically related word precedes the target. If the base word mediates the recognition of its derivatives, the prior processing of such words should speed up the recognition of the base. Repetition priming effects have been reported for English. Stanners *et al.* (1979) found that the facilitatory effect of regularly inflected verb forms on their stems was as large as the repetition effect of the stem on itself: *walks* primed *walk* to the same extent as *walk* primed *walk*. Similar results are reported for English derivation by Stanners, Neiser and Peinton (1979) and Forster and Davies (1984). These findings are not undisputed, however. Fowler, Napps and Feldman (1985) present evidence that repetition priming effects are episodic rather than lexical (see Henderson, 1989, for a review).

The role of morphology at the central level can also be studied by manipulating word frequency. If the recognition of a complex word involves prior access to its stem representation, and if this stem representation is sensitive to frequency of use, then the summed frequencies of all words containing this stem

should be a more reliable predictor of response latencies than the frequency of the surface form itself. Taft (1979) reported faster lexical decision times for words with higher stem frequency and the same surface frequency. When matched on stem frequency, an effect of word frequency remained. These results have been partially replicated by Colé, Beauvillain and Segui (1989) who observed a cumulative frequency effect for derivationally suffixed words in French, but not for prefixed words. Similarly, Bradley (1979) found a cumulative frequency effect for the phonologically transparent English suffixes *-er*, *-ment* and *-ness*, but not for phonologically far less regular *-ion*. Related studies by Lukatela *et al.* (1980) suggest that in Serbo-Croatian, oblique forms are more dependent on cumulative frequency effects than nominative base forms. In sum, there is some, be it not unequivocal, evidence that at least some complex word types require access to a central stem representation.

An experimental technique that sheds light on morphological effects at the level of access representations is phoneme monitoring. This task requires subjects to monitor for a given phoneme and to press a button as soon as this phoneme has been heard. Frauenfelder (personal communication) used this technique to study the speed with which morphemes in Turkish words become available. He obtained evidence that both roots and affixes are immediately accessed as independent units, suggesting that access representations for these roots and affixes are available for the activation of the corresponding central representations.

Additional evidence can be obtained by manipulating the morphological status of nonwords. Laudanna, Burani and Cermele (1994) compared reaction times to nonwords with different types of prefix. One dimension along which prefixes differ is the number of pseudo-prefixed words, the words that appear to contain that prefix, but where no prefix is actually involved (as for instance *re-* in English *reindeer*). As pseudowords by definition lack access and central representations, any effect of a potential morpheme entails the existence of an access as well as a central representation for that morpheme. What they found was that pseudowords containing prefixes with many pseudo-prefixed words in the language required less time to reject than pseudowords with prefixes with few such pseudo-prefixed words. These results suggest that segmentation will not be carried out if there is a high chance of obtaining erroneous results, to our mind a desirable property of the processing system. On the other hand, Taft and Forster (1975) observed longer response latencies for existing words containing a pseudo-prefix compared to controls without pseudo-prefixes, which they take to imply obligatory parsing of all potential prefixes. We will return to this issue below.

Licensing

An important study that has explicitly addressed the issue of licensing in word recognition is that by Caramazza, Laudanna and Romani (1988). These authors explored the role of licensing by presenting subjects with inflected pseudowords of various kinds in a lexical decision task. They compared the reaction times to pseudowords with either an existing stem and a pseudo-affix, a non-existing stem and a real affix, and the combination of a real stem and a real affix that is from an incorrect declension class. Typically, the longest rejection times were obtained for the latter kinds of stimulus. Correct no responses crucially require access to the relevant subcategorization information at a deeper level of processing than simple morpheme identification.

Composition

Relatively few studies have addressed the composition process, that is, the computation of meaning. The existing studies focus on the processing of novel compounds only. Coolen, Van Jaarsveld and Schreuder (1991) obtained evidence suggesting that the semantic interpretation of novel compounds is an immediate and automatic process. This conclusion was based on the observation that it is more difficult for subjects to recognize a compound as novel when that compound has an obvious interpretation. Such interference effects are generally assumed to indicate automatic processing. Furthermore, Coolen, Van Jaarsveld and Schreuder (1991) claim that the combination process is autonomous in the sense that the meaning of a novel compound is constructed from the meanings of its parts in interaction with general world-knowledge rather than by means of analogy-based operations on existing compounds in the lexicon. Although these results show that novel words can be interpreted quickly and efficiently, the exact mechanisms which regulate the interpretation remain unknown.

Psycholinguistic Models

In this section we discuss a number of psycholinguistic models for morphological processing. Most of these models have not been implemented in any way, with the exception of Gasser (1994a, 1994b), whose connectionist model of affix acquisition is a real, operational computational model. The latter model is discussed later, the former models are reviewed below.

Verbal Models

The available experimental evidence for morphological processing is, unfortunately, too weak to sufficiently constrain the modelling effort. Hence it is possible for some models of word recognition not to assign any relevance to morphological structure (see e.g., Butterworth, 1983; Seidenberg, 1988), whereas others assume obligatory parsing for every complex word. Here we focus on those models that assume at least some form of morphological processing.

The first model to take morphological structure into account is the prefix-stripping model proposed by Taft and Forster (1975). They hypothesized that prefixes are removed prior to lexical access. The remaining string is used as the access code for a serial search in a so-called access bin. The matching entry in this bin contains a pointer to the list of possible full matches in the master lexicon. This schema was proposed as a way of increasing the efficiency of the serial search for prefixed words. Experiments testing the predictions of this model have led to a series of contradictory results. Even worse, it has been shown that the model is rather unattractive computationally. In fact, the efficiency of a serial search model decreases by roughly a factor of two when a prefix stripping strategy is incorporated, due to the large numbers of pseudo-prefixed word tokens in daily language (from 50 per cent up to 95 per cent for English, depending on one's definition of pseudo-prefixation, see Schreuder and Baayen, 1994).

A more interesting verbal model is the so-called Augmented Addressed Morphology Model (AAM) developed by Burani and Caramazza (1987), Caramazza, Laudanna and Romani (1988), Burani and Laudanna (1993). This model claims that the segmentation phase is generally bypassed in word recognition. All known words, whether morphologically complex or not, have their own access representations in this model. The (modality-specific) access representation of a complex word is linked up to the modality-free central representations of its constituent morphemes. In this central lexicon the representation of the complex whole is computed on-line, although the possibility is left open that very high-frequency complex words may acquire full lexical representations of their own. In order to handle complex neologisms, it is assumed that affixes have access representations in addition to lexical representations. The access route that exploits the full access representations is stipulated to be always faster than the segmentation route. It will be clear that this is a highly underspecified process model. Although Caramazza, Laudanna and Romani are well aware of the importance of phonological and semantic transparency, they have not spelled out what kinds of computation could be involved. The model in its present form is as yet not sufficiently detailed to allow a computational implementation.

A closely related model is the Morphological Race Model (MRM) formulated by Frauenfelder and Schreuder (1992). As in the AAM, two access routes are proposed. Unlike in the AAM, the outcome of the race is not fixed beforehand. In the MRM, both routes operate in parallel, so that it is not excluded that known complex words are nevertheless recognized by the morphological route, that is, by segmentation and subsequent composition. In addition, these authors postulate a mechanism that strengthens the access representations of the constituent(s) that contributed to the winning route. In the case that the direct route is the first to complete access, the resting activation level of the full form access representation of the word is slightly increased. If the parsing route completes first, the resting activation levels of stem and affix access representations are increased. Semantically non-compositional words can be accessed only via the whole-word route. Hence, only the access representations of affixes that participate in many semantically transparent words will obtain high resting activation levels. This is the way in which differences in productivity of affixes are accounted for. The high activation levels of productive affixes enable the system to process both novel complex words and relatively low-frequency transparent words that have been encountered before, but not often enough to allow the whole-word address route to win. However, the MRM is a verbal model that is too unspecified to allow computational modelling.

An extension of the MRM that addresses the role of semantic transparency and its effect on the activation levels of access representations in a more principled way can be found in Schreuder and Baayen (1995). Each of the three recognition stages, segmentation, licensing, and composition, is modelled in some detail using an architecture that combines spreading activation mechanisms with a symbolic parser operating on representations that have reached sufficiently high activation levels. The segmentation of the input into its constituents is accomplished by the spreading activation mechanisms. Licensing and composition are carried out by a unification-based symbolic parser. It is a dual route model in which the routes interactively converge on the correct meaning. The model has three representational levels: the level of access representations, a level of concept nodes, and a level of syntactic and semantic nodes. Activation is allowed to flow back from the semantic and syntactic nodes via the concept nodes to the access representations. This mechanism ensures that semantic transparency is a driving factor in determining the role affixes and stems play in the recognition of morphologically complex words. An *ad hoc* mechanism providing positive feedback for the winning route as in the MRM is no longer necessary. The importance of transparency for the acquisition of affixes and the consequences of differences in age of acquisition for the adult processing system are also stressed.

An Implemented Computational Model

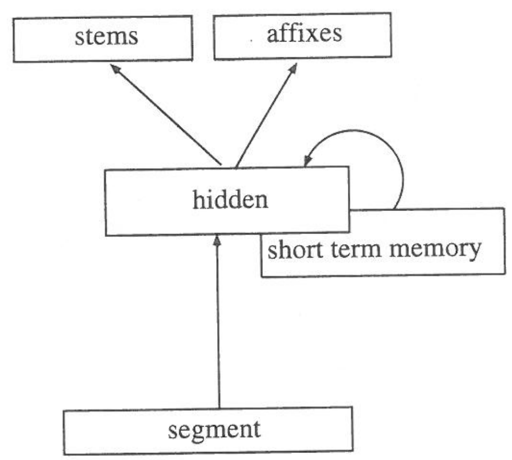
The acquisition of the skill to map form onto meaning for morphologically complex words is studied in detail by Gasser (1994a, 1994b), using connectionist architectures. Interestingly, Gasser does not focus on the issue of whether a single network can accommodate both regular and irregular morphology at the same

time, a hotly debated issue in the acquisition literature (Rumelhart and McClelland, 1986; MacWhinney and Leinbach, 1991; Pinker, 1991). Instead, his interest lies with which network architectures are required for the learning of the mapping of form onto meaning in the auditory modality for a wide variety of regular morphological processes (suffixation, infixation, prefixation, circumfixation, deletion, vocalic alternation, and reduplication), and what difficulties specific properties of these morphological processes pose for learning.

Following Norris (1990), Gasser makes use of a recurrent network with three layers to model spoken word recognition. The first layer encodes single phonemes as they become available in time. This phoneme layer is connected to a hidden layer. At each time step, the contents of this hidden layer are copied into a second hidden layer that serves as a short-term memory. At the next time step, the first hidden layer receives input from both the phonological layer and this second hidden layer. In this way the network is enabled to 'remember' the phonemes that were presented to the network at earlier time steps. Finally, the first hidden layer is connected to a semantic layer, where each node represents a full concept. The semantic layer contains two separate sets of units, one set representing the meanings of content words (labelled *stems* in the slightly simplified architecture shown in Figure 7.6), the other the 'grammatical meanings' of affixes (labeled *affixes*). In most of his experiments, Gasser trained the network on a set of 30 roots from a controlled artificial (but linguistically well-motivated) language. Each experiment studied the behaviour of the network for one particular morphological process (prefixation, suffixation, infixation, etc.) The network was trained on 2/3 of the possible combinations. Its performance was evaluated on the remaining — novel — combinations. The networks were trained using conventional backpropagation.

In a first series of experiments, Gasser tested the functionality of the network of Figure 7.6. What he found was that the network has severe difficulties in

Figure 7.6: Gasser's (1994a, 1994b) non-modular connectionist network

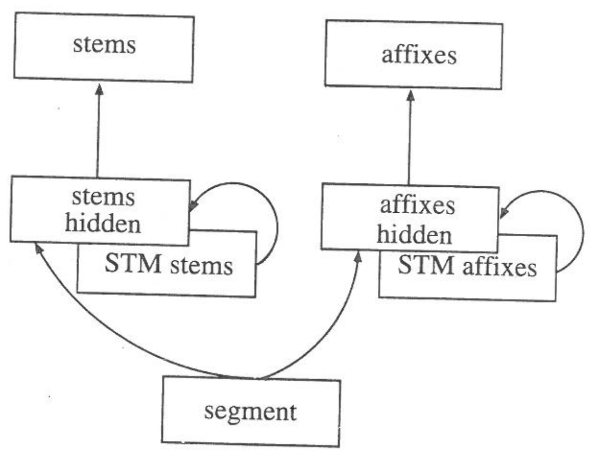


recognizing both the stem and the affix of novel complex words simultaneously. For instance, it performs moderately well for stem recognition for suffixed words, but has great difficulties in getting the suffix right. Conversely, it has no difficulty in recognizing prefixes, but for prefixed words recognition of the stem is abysmal. Apparently, the network has severe problems due to the hidden layer being required to perform two separate tasks at the same time: affix recognition and stem recognition. To avoid overtaxing the hidden layer, Gasser studied a modular network in which stems and affixes have their own hidden layer, as shown in Figure 7.7. Performance for affix recognition improved to nearly 100 per cent for a wide variety of morphological operations, infixation and vocalic alternation being the exceptions. Stem recognition also increased, be it to a much lower level of accuracy (40 per cent–75 per cent). In the case of circumfixation, its performance on root recognition was very poor. Apparently, stem recognition is severely impaired when both the beginning and the end of the stem are 'masked' by affixed segments.

In the system shown in Figure 7.7, the learning of affixes is facilitated by the pre-wiring of the modular structure of the network: the hidden layers are already specialized recognition devices for stems and affixes respectively. Gasser shows that it is, to some extent, possible to let this modular organization arise as part of the learning process. He first trained the net on the recognition of stems. Following this, he added a number of so-called gating units monitoring the connections from the two sets of hidden layers to both stem and affix units. He observed that the network tends to allocate stem recognition to one set of hidden units, and suffix recognition to the other. However, when the network is presented with both prefixes and suffixes, hardly any specialization takes place.

Finally, Gasser shows that the system shown in Figure 7.7 cannot learn rules involving reduplication. He argues that reduplication can be handled only by extending the system with a hidden layer that is geared to recognizing syllables.

Figure 7.7: Gasser's (1994a, 1994b) modular connectionist network



That the recognition system must be allowed to exploit prosodic information is a conclusion that is shared by Sproat, who, in the symbolic domain, shows that a KIMMO-like system by itself cannot adequately handle the complex reduplication patterns of a language such as Warlpiri (Sproat, 1992, pp. 164–70).

It will be clear that the modelling of the acquisition of morphological rules is far from trivial. Gasser's networks appear to be reasonably effective for the learning of a single morphological operation. It is unclear how well this architecture will be able to deal with words with multiple affixes such as *ondenkbaarheid*, 'un-think-able-ness', where on- is not subcategorized for attaching to the verb stem denk, and where the parsing *[[on + [denk + baar]] + heid]* should be selected and not the possible parsing *[on + [[denk + baar] + heid]]*. In addition, it should be noted that languages are often characterized by large numbers of morphologically complex but semantically non-compositional words. If one's goal is to model the acquisition of the full vocabulary and lexical rules of a language, the presence of large numbers of words with affixes that do not allow a regular mapping onto semantics may well cause serious problems for modular networks of this kind, especially in the light of the strong correlation between higher token frequencies and higher degrees of semantic opacity. A number of verb-forming prefixes in Dutch, for instance, are most often encountered in semi-transparent and opaque formations (Schreuder and Baayen, 1994, 1995, see also Lieber and Baayen, 1994). If a learning system is to master the transparent semantics of such affixes, it must do so in an environment that is characterized by a very low signal-to-noise ratio. The large numbers of pseudo-prefixed words observed for English and Dutch lower this signal-to-noise ratio even more. It is as yet unclear which systems that make use of statistical learning can handle distributions of this kind. Nevertheless, Gasser's work is an impressive first step in the computational modelling of morphological processing.

Discussion

The linguistic computational models discussed are clearly superior to the psycholinguistic models with respect to the range and depth of the linguistic phenomena they cover. Furthermore, compared with Gasser's (1994a, 1994b) model, they perform far better. Nevertheless, in their present form these models are unsatisfactory from a psycholinguistic point of view, for a number of reasons. First consider KIMMO. Koskenniemi (1983, p. 134–6) claims that his model provides a psychologically valid description of morphological processing. However, there are several aspects of this approach that are unsatisfactory. First, in contrast to Gasser's model, KIMMO is not a learning system. All morphological rules are embodied in the transducers once and for all. The question of how a system might acquire such — often intricately interacting — transducers is not answered. Second, the L-tree mechanism is an elegant technical solution for lexical look-up, but it is psycholinguistically implausible. The use of an L-tree predicts that certain classes of nonword should be rejected in much less time than it takes to respond to real words, contrary to fact (Forster, 1989). In addition, word frequency and cumulative frequency effects do not emerge naturally in systems based on the L-tree format. Finally, if lexical access in the auditory modality proceeds on the basis of metrical segmentation strategies as suggested by Cutler and Norris (1988), the idea of a simple left-to-right scan through the word may well be too simplistic.

Next consider Ritchie *et al.*'s (1992) morphological parser for English. An attractive feature of this model is that the number of different morphological rules embodied in the system are kept at a minimum. The main computational burden lies with the computationally relatively simple unification of the lexical representations associated with the morphemes. On the negative side, the model in its present form is designed to parse each and every input string. This property may be desirable from a linguistic (and a practical) point of view, but it is psychologically unsatisfactory, especially for languages such as English where conclusive evidence for an obligatory on-line role for morphological structure has not been obtained. In languages with simple morphologies, high-frequency complex words may very well be recognized using a form of direct access. Furthermore, one would not expect the constituents of semantically opaque words to play a role in production or perception. Finally, note that in this computational approach the way in which the system discovers morphemes and their properties is again left unspecified.

Summing up, it is clear that compared to many of the computational models discussed in the present volume the available psycholinguistic models of morphological processing are underspecified and rudimentary. No operational models predicting reaction times are available. Undoubtedly, this state of affairs can be traced to the complexity of the issues involved. For psycholinguistics, the challenge of building an operational model of morphological processing has yet to be met.

References

ANDERSON, S. R. (1992) *A-Morphous Morphology*, Cambridge: Cambridge University Press.

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

ARONOFF, M. (1994) *Morphology by Itself: Stems and Inflectional Classes*, Cambridge, MA: MIT Press.

BEARD, R. (1992) *Lexeme-Morpheme Base Morphology. A General Theory of Inflection and Word Formation*, Unpublished manuscript.

BRADLEY, D. (1979) 'Lexical representation of derivational relation', in ARONOFF, M. and KEAN, M. L. (Eds) *Juncture*, Sarasota: Anma Libri, pp. 37–55.

BURANI, C. and CARAMAZZA, A. (1987) 'Representation and processing of derived words', *Language and Cognitive Processes*, 2, pp. 217–27.

BURANI, C. and LAUDANNA, A. (1993) 'Units of representation for derived words in the lexicon', in FROST, R. and KATZ, L. (Eds) *Orthography, Phonology, Morphology, and Meaning*, Amsterdam: Elsevier/North-Holland.

BUTTERWORTH, B. (1983) 'Lexical representation', in BUTTERWORTH, B. (Ed) *Language Production: Vol 2. Development, Writing and Other Language Processes*, London: Academic Press, pp. 257–94.

CARAMAZZA, A., LAUDANNA, A. and ROMANI, C. (1988) 'Lexical access and inflectional morphology', *Cognition*, 28, pp. 297–332.

COLÉ, P., BEAUVILLAIN, C. and SEGUI, J. (1989) 'On the representation and processing of prefixed and suffixed derived words: A differential frequency effect', *Journal of Memory and Language*, 28, pp. 1–13.

COOLEN, R., VAN JAARSVELD, H. and SCHREUDER, R. (1991) 'The interpretation of isolated novel nominal compounds', *Memory and Cognition*, 19, pp. 341–52.

CUTLER, A. and NORRIS, D. G. (1988) 'The role of strong syllables in segmentation for lexical access', *Journal of Experimental Psychology: Human Perception and Performance*, 14, pp. 113–21.

FORSTER, K. I. (1989) 'Basic issues in lexical processing', in MARSLEN-WILSON, W. (Ed) *Lexical Representation and Process*, Cambridge, MA: MIT Press, pp. 75–108.

FORSTER, K. I. and DAVIS, C. (1984) 'Repetition priming and frequency attenuation in lexical access', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, pp. 680–98.

FOWLER, C. A., NAPPS, S. E. and FELDMAN, L. B. (1985) 'Relations among regularly and irregularly morphologically related words in the lexicon as revealed by repetition priming', *Memory and Cognition*, 13, pp. 241–55.

FRAUENFELDER, U. H. and SCHREUDER, R. (1992) 'Constraining psycholinguistic models of morphological processing and representation: The role of productivity', in BOOIJ, G. E. and VAN MARLE, J. (Eds) *Yearbook of Morphology 1991*, Dordrecht: Kluwer, pp. 165–83.

GASSER, M. (1994a) 'Modularity in a connectionist model of the acquisition of morphology', in *Proceedings of the 15th International Conference on Computational Linguistics (COLING '94), Kyoto*, pp. 214–20.

GASSER, M. (1994b) 'Acquiring receptive morphology: A connectionist model', in *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics, Las Cruces*, pp. 279–86.

GESENIUS, W. and KAUTZSCH, E. (1909) *Hebräische Grammatik*, Leipzig: Vogel.

GLEASON, H. A. (1955) *Workbook in Descriptive Linguistics*, New York: Holt, Rinehart & Winston.

HANKAMER, J. (1989) 'Morphological parsing and the lexicon', in MARSLEN-WILSON, W. (Ed) *Lexical Representation and Process*, Cambridge, MA: MIT Press, pp. 392–408.

HEEMSKERK, J. (1993) 'A probabilistic context-free grammar for disambiguation in morphological parsing', in *Proceedings of the 6th Conference of the European Chapter of the Association for Computational Linguistics, Utrecht*, pp. 183–92.

HEEMSKERK, J. and VAN HEUVEN, V. (1993) 'MORPA: a morpheme lexicon-based morphological parser', in VAN HEUVEN, V. and POLS, L. (Eds) *Analysis and Synthesis of Speech, Strategic Research Towards High-Quality Text-to-Speech Generation*, Berlin: Mouton/De Gruyter.

HENDERSON, L. (1989) 'On mental representation of morphology and its diagnosis by measures of visual access speed', in MARSLEN-WILSON, W. (Ed) *Lexical Representation and Process*, Cambridge, MA: MIT Press, pp. 357–91.

JACKENDOFF, R. (1991) *Semantic Structures*, Cambridge. MA: MIT Press.

JENSEN, J. T. (1990) *Morphology: Word Structure in Generative Grammar*, Amsterdam: John Benjamins.

KOSKENNIEMI, K. (1983) *Two-Level Morphology. A General Computational Model for Word-Form Recognition and Production*, Unpublished doctoral dissertation, University of Helsinki.
KOSKENNIEMI, K. (1984) 'A general computational model for word-form recognition and production', in *Proceedings of the 10th International Conference on Computational Linguistics* (*COLING '84*), pp. 178–81.
KNUTH, D. E. (1973) *The Art of Computer Programming*, Reading, MA: Addison-Wesley.
LAUDANNA, A., BURANI, C. and CERMELE, A. (1994) 'Prefixes as processing units', *Language and Cognitive Processes*, 9, pp. 295–316.
LIEBER, R. (1980) *On the Organization of the Lexicon*, Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.
LIEBER, R. (1987) *An Integrated Theory of Autosegmental Processes*, Albany: State University of New York Press.
LIEBER, R. (1992) *Deconstructing Morphology. Word Formation in Syntactic Theory*, Chicago: University of Chicago Press.
LIEBER, R. and BAAYEN, R. H. (1994) 'Verbal prefixes in Dutch: A study in lexical conceptual structure', in BOOIJ, G. E. and VAN MARLE, J. (Eds) *Yearbook of Morphology 1993*, Dordrecht: Kluwer, pp. 51–78.
LUKATELA, G., GLIGORIJEVIC, B., KOSTIC, A. and TURVEY, M. T. (1980) 'Representation of inflected nouns in the internal lexicon', *Memory and Cognition*, 8, pp. 415–23.
MACWHINNEY, B. and LEINBACH, J. (1991) 'Implementations are not conceptualizations: Revising the verb learning model', *Cognition*, 40, pp. 121–57.
MARTIN, W. (1988) 'Subtractive morphology as dissociation', in *Proceedings of the Seventh West Coast Conference on Formal Linguistics*, Stanford: Linguistic Association.
MCCARTHY, J. (1982) *Formal Problems in Semitic Morphology and Phonology* (Doctoral dissertation, Massachusetts Institute of Technology, 1979). Indiana University Linguistics Club.
MCQUEEN, J. M. and CUTLER, A. (in press) 'Morphology in word recognition', in ZWICKY, A. M. and SPENCER, A. (Eds) *The Handbook of Morphology*, Oxford: Blackwell.
NORRIS, D. (1990) 'A dynamic-net model of human speech recognition', in ALTMANN, G. T. M. (Ed) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*, Cambridge, MA: MIT Press, pp. 87–104.
PINKER, S. (1991) 'Rules of language', *Science*, 253, pp. 530–5.
RITCHIE, G. D., PULMAN, S. G., BLACK, A. W. and RUSSELL, G. J. (1987) 'A computational framework for lexical description', *Computational Linguistics*, 13, pp. 290–307.
RITCHIE, G. D., RUSSEL, G. J., BLACK, A. W. and PULMAN, S. G. (1992) *Computational Morphology. Practical Mechanisms for the English Lexicon*, Cambridge, MA: MIT Press.
RUMELHART, D. E. and MCCLELLAND, J. L. (1986) 'On learning the past tenses of English verbs', in MCCLELLAND, J. L., RUMELHART, D. E. and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Vol. 2. Psychological and Biological Models*, Cambridge, MA: MIT Press, pp. 216–71.
SCHACHTER, P. and OTANES, F. (1972) *Tagalog Reference Grammar*, Berkeley: University of California Press.
SCHANE, S. A. (1973) *Generative Phonology*, Englewood Cliffs: Prentice-Hall.
SCHREUDER, R. and BAAYEN, R. H. (1994) 'Prefix-stripping re-revisited', *Journal of Memory and Language*, 33, pp. 357–75.
SCHREUDER, R. and BAAYEN, R. H. (1995) 'Modelling morphological processing', in FELDMAN, L. (Ed) *Morphological Aspects of Language Processing*, Hillsdale, NJ: Lawrence Erlbaum, pp. 131–54.
SEIDENBERG, M. S. (1988) 'Reading complex words', in CARLSON, S. N. and TANENHAUS, M. K. (Eds) *Linguistic Structure in Language*, Dordrecht: Kluwer, pp. 53–106.
SELKIRK, E. O. (1980) *The Syntax of Words*, Cambridge, MA: MIT Press.

SHIEBER, S. M. (1986) *An Introduction to Unification-Based Approaches to Grammar*, Stanford: Center for the Study of Language and Information.
SPROAT, R. (1992) *Morphology and Computation*, Cambridge, MA: MIT Press.
STANNERS, R. F., NEISER, J. J., HERNON, W. P. and HALL, R. (1979) 'Memory representation for morphologically related words', *Journal of Verbal Learning and Verbal Behavior*, 18, pp. 399–412.
STANNERS, R. F., NEISER, J. J. and PEINTON, S. (1979) 'Memory representation for prefixed words', *Journal of Verbal Learning and Verbal Behavior*, 18, pp. 733–43.
TAFT, M. (1979) 'Recognition of affixed words and the word frequency effect', *Memory and Cognition*, 7, pp. 263–72.
TAFT, M. and FORSTER, K. I. (1975) 'Lexical storage and retrieval of prefixed words', *Journal of Verbal Learning and Verbal Behavior*, 14, pp. 638–47.