

Some studies of lexical access at the Max Planck Institute for Psycholinguistics

Willem J. Levelt

Max Planck Institute for Psycholinguistics, Nijmegen

1. Preliminaries: The Max Planck Institute for Psycholinguistics. The Max Planck Institute for Psycholinguistics was founded in 1980, after a three-year try-out period as a Max Planck Projectgroup. It is one of 53 institutes for basic research operating under the wings of the Max Planck Society. The Max Planck Society is an independent West German scientific organization which promotes basic research for the benefit of the public. It is largely financed by public funds made available jointly by the Federal and State Governments.

The Society's postwar history is one of intensive European cooperation. It is an important participant in such institutions as the European Space Agency (ESA), the European Nuclear Research Organization (CERN), the Institute for Radioastronomy in the Millimeter Wavelength Range (IRAM), and other research endeavors outside the German Federal Republic. There is also a Max Planck art history institute, the Bibliotheca Hertziana in Rome. The Society, moreover, has an extensive system of research grants for non-Germans.

The establishment of a Max Planck Institute in the Netherlands was a further move towards 'europeanization' of the Society. Nijmegen University, a few kilometers from the German border, already had a strong research tradition in psycholinguistics. The university could provide an excellent sounding board for the Society's new initiative. It could, in addition, provide expertise and cooperation in neighboring areas, such as neurology and artificial intelligence.

The Society's gesture was met with enthusiasm on the Dutch side. Nijmegen University established a peer group, the Interfaculty Research Unit for Language and Speech, presently in full bloom. The Netherlands Organization for Scientific Research (NWO) has provided substantial additional funds since the establishment of the Projectgroup. These funds are presently supporting a cooperative research project between the Institute and the Interfaculty Research Unit on aphasia in adults. Also, the Institute became the coordinating center for a large-scale research project of the European Science Foundation, called 'Second Language Acquisition in Adults', a study of natural language acquisition by foreign workers. In this project, which was carried out in five European countries simultaneously, young adults from seven different mother tongues were followed longitudinally over a

period of 30 months as they acquired their host language. The acquisition of Dutch in this project was studied by a research team at Tilburg University under the direction of Guus Extra (see his paper in the present volume). This ESF project was funded through the ESF by a large variety of European research councils, among them NWO, as well as by the Max Planck Society.

In the Max Planck Institute psychologists and linguists cooperate in three areas of interdisciplinary research--the acquisition of language, the comprehension of language, and the production of language. Given the size of the institute (there are presently 90 people under some form of contract, among them about 35 scientists), it is impracticable to fully review its research here. I rather refer to our annual reports, which appear yearly around April 30. Instead, I will discuss a few studies of lexical access with which I have recently been involved.

2. Lexical access in speaking and speech comprehension. One of the most fascinating problems in psycholinguistics is the language user's skillful accessing of lexical items. In speaking and listening, accessing rates of two to three words per second are normal, but six to seven words per second are not exceptional. And the number of accessing failures in speaking and listening is remarkably low. The issue then is: What kinds of mechanism allow the language user to access words at these high speeds and with this degree of accuracy in a mental lexicon which contains several tens of thousands of items?

One can distinguish two stages in the accessing process. The first one can be called 'convergence', the second one 'unpacking'. For the speaker, convergence consists of zooming in on the one (if any) lexical item which appropriately expresses the intended concept in the given context of discourse. It is a semantically driven process. Unpacking is the retrieval of that item's phonological features and their use in preparing the articulatory program. I will shortly return to some of our experimental studies of semantic convergence and phonological unpacking in the production of speech.

For the hearer, convergence is the process by which the signal input plus context leads to the activation and selection of the 'correct' lexical item in the mental lexicon. Unpacking is the making available of the selected item's semantic make-up, and its integration into the interpretation of the sentence and the larger discourse. The Institute's research is, specifically, concerned with the question how 'bottom-up' signal information does or does not interact with 'top-down' contextual information in the course of selecting the item. The theory developed in this connection is known as the 'cohort theory' (Marslen-Wilson and Tyler 1980, 1981). Before turning to issues of access in production, I will say a few words about this cohort theory. For the issues in production are formally very similar to those in comprehension, whereas we have a better understanding of what is going on in comprehension than in production. The cohort theory can provide us with some guiding hypotheses.

3. Lexical access in comprehension. Some remarks on the cohort theory. Let us first consider the recognition of single words, spoken in isolation. The 'classical' cohort model for this case is, in summary, as follows. Some initial part of the spoken word, i.e. the first 100 to 150 milliseconds of the speech

signal, is used by the listener to establish contact with a set of lexical candidates. They are just those items in the mental lexicon which share that word-initial part. So, for instance, when the word *captain* is spoken, the initial part of the signal, *cap*, suffices to activate not only the lexical item *captain*, but also items such as *capsule*, *capital*, and so on. This set of lexical items is called the 'word initial cohort'.

As more sensory input is acquired, the cohort is correspondingly reduced, until a single candidate compatible with the input signal remains. At the moment that *capt* has been delivered, alternatives such as *capsule* and *capital* disappear. But a few others still survive, such as *captor* and *captive*. The word is identified as soon as it is the only survivor in this process, i.e., as soon as it can be distinguished from all other activated word candidates. This is usually the case before the end of the word. For *captain* it is the case during its last vowel. As soon as the word has been identified, it can be interpreted or, as we said, semantically 'unpacked'.

There is, at present, good experimental evidence for the following two essential aspects of the model: (1) the lexical alternatives to the target word, i.e., the elements of the word-initial cohort, are indeed activated in the listener's mind upon presentation of the word-initial sequence; (2) the isolated word is indeed identified as soon as the incoming signal makes it uniquely distinguishable from all other candidates (cf. Marslen-Wilson 1987, Zwitserlood 1985, 1989).

So far the model is compatible with the view that the phase of (phonetic/phonological) convergence precedes the phase of (semantic) 'unpacking' or integration. But the picture changes when we consider the recognition of words in spoken context. To elaborate the example, consider the following text (after Zwitserlood 1985):

With dampened spirits the sailors stood around the grave. They mourned the loss of their CAPTAIN.

A candidate such as *captain*, which is a plausible continuation in the context, is identified earlier than one which is not a plausible continuation (such as *captor*). In other words, the word-initial cohort can be reduced more rapidly when there is semantic and/or syntactic supporting context. It turns out, however (Zwitserlood 1985, 1989), that the implausible alternatives (such as *captor* in the example) do become activated, just as in the isolated word condition discussed above. The inevitable conclusion is that the phases of convergence and of unpacking overlap in speech comprehension, and that the two processes interact: convergence is speeded up by semantic interpretation. Let us now turn to lexical access in production and discuss some very similar questions.

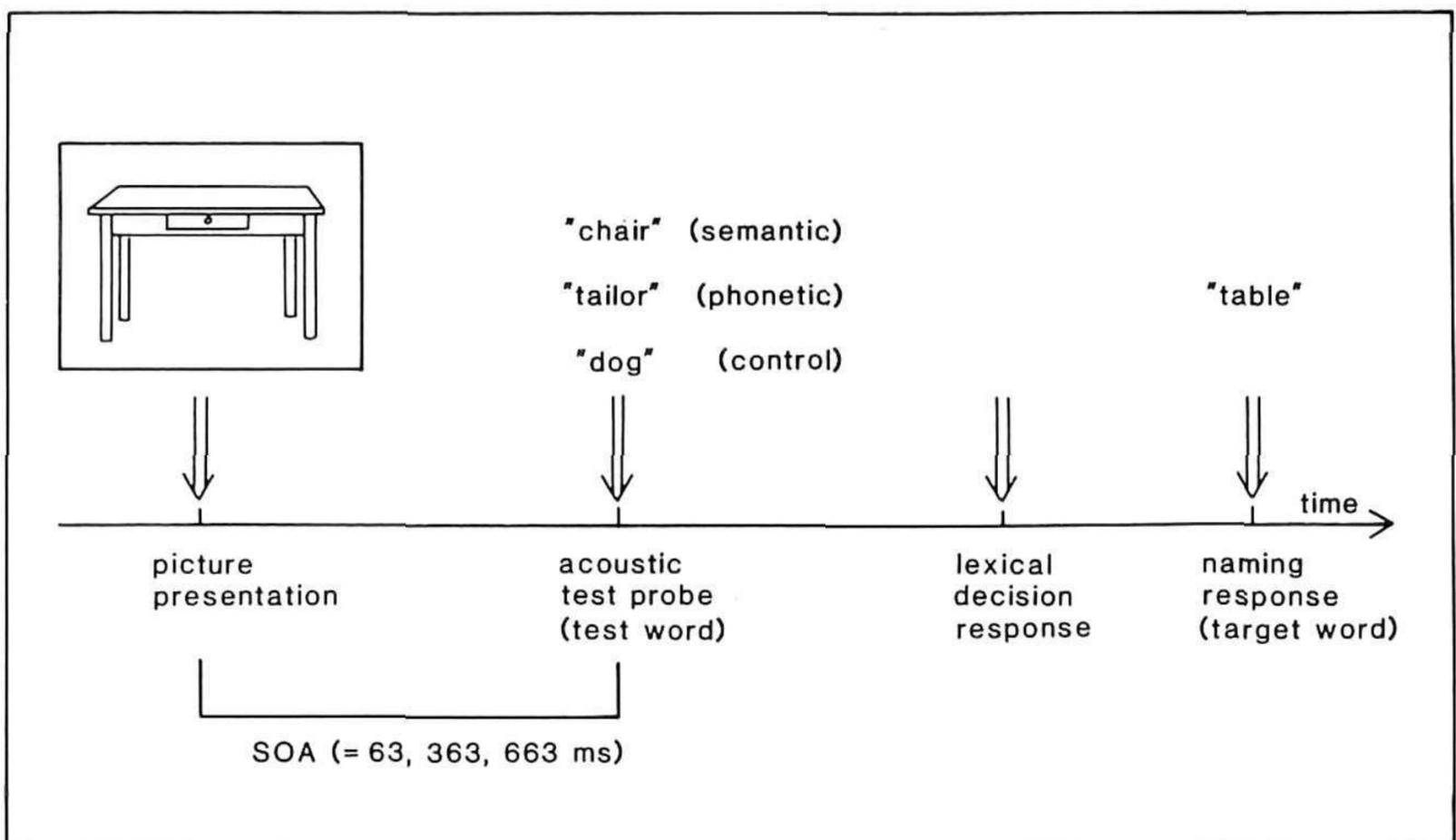
4. Is lexical access in production a two-phase process? Since the modern resumption of speech error research, it has been argued that speakers access words in two successive phases (see, for instance, Fromkin 1971). During the first phase they converge on what is now commonly called a *lemma*, an item in a semantic/syntactic lexicon. Each lemma points to an address in a phonological lexicon, where the word form can be found. During the second

phase the word form is accessed and 'unpacked' in order to generate a phonetic program for the articulation of the word. For reviews of this two-phase theory, see Butterworth (forthcoming) and Levelt (1989). The well-known tip-of-the-tongue phenomenon (Brown and McNeill 1966) arises, according to this theory, when the speaker successfully completes the first phase of access, but blocks on the second one.

In a recent series of experiments, Herbert Schiefers, Antje Meyer, and I (cf. Meyer 1988) have traced the time course of these two phases, in an effort to determine whether they are or are not overlapping. The paradigm was one of object naming, where subjects were asked to name, one after another, pictures presented to them. The situation is, therefore, comparable to the isolated word condition discussed above: lexical access without supporting context.

The rather complicated experimental procedure is exemplified in Figure 1. Subjects named objects presented to them, but occasionally they were given a secondary task. At certain trials an acoustic test probe was presented shortly after presentation of the picture. This test probe could be a normal word or a nonword. Every time such a test probe was presented, the subject was to push a 'yes' or a 'no' button; 'yes' if the test probe was a word and 'no' otherwise. This task is called 'lexical decision'. For all cases to be considered here the test probe was a word, and the subject (normally) pushed the 'yes' button. Our experimental measure was the subject's push response latency from the onset of the test probe, the so-called lexical decision latency.

Figure 1. A paradigm for lexical decision during object naming.



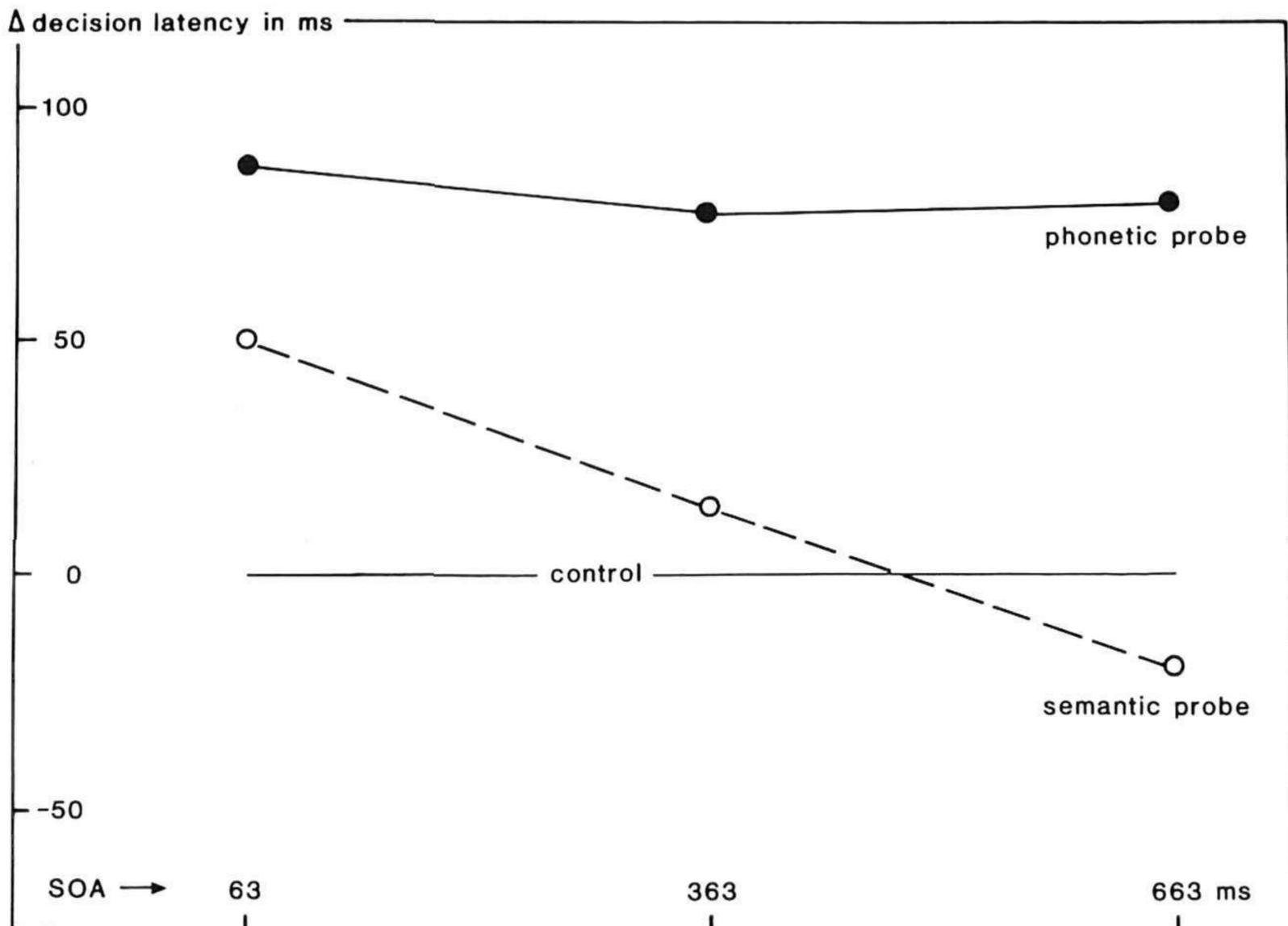
We used different kinds of test probes in order to measure semantic and phonological activation. If, for instance, the object to be named was a table so that the speaker was in the course of preparing the target word *table*, we could unexpectedly present as acoustic test probe the word *chair*. This test word is semantically related to the target word. We expected that the lexical

decision latency would be affected when a semantically related probe appeared in the first, semantic phase of lexical access. Similarly, we could give *tailor* as an acoustic probe. This test word is phonetically related to the target word *table*. We expected a response latency effect if such a test probe fell in the second, phonological phase of access. We also had various controls, such as an unrelated test word. The word *dog*, for instance, is neither semantically nor phonologically related to the target word *table*. Such unrelated test words were used to establish base line values for the lexical decision latencies in the experiment.

The test word could be presented at one of three moments ('stimulus onset asynchronies', or SOAs) after presentation of the picture that was to be named. We expected evidence for semantic activation at an early moment, right after presentation of the picture (at an average SOA of 63 milliseconds), and evidence for phonological activation at a late moment, shortly before the naming response (with average SOA of 663 milliseconds). We also added an intermediate SOA condition, with an average SOA of 363 milliseconds.

Our findings confirmed these expectations. A summary of some main results is presented in Figure 2.

Figure 2. Semantic and phonetic interference for different SOAs.



The figure shows the increase of lexical decision latency in going from the control probes (like *dog*) to the semantic or phonetic probes (like *chair* or *tailor*). For the late SOA (663 ms.) there is a significant delay for the reaction to the phonetic probe. This means that the available phonological form of the target word (*table* in the example) interferes with the subject's lexical decision.

There is, however, no late effect on the semantic probe (like *chair*), which indicates that the speaker is really in the second phase of lexical access, the unpacking of phonological form. As expected, a semantic effect is found for the short SOA condition (63 ms.); the response to a semantically related probe word (*chair* in the example) is significantly delayed when the probe is given right after the picture. This indicates that the speaker is then in the first, semantic phase of lexical access.

But there is an important additional result: there is also a significant effect for the *phonetic* probe at this early moment. Does this indicate that there is overlap between the two phases of access? Does it, in particular, mean that activation of the lemma immediately results in activation of the phonological form it points to? So far the results tell us that phonological activation of a word can survive its semantic or lemma activation. But can a lemma be activated without immediate coactivation of the phonological form it points to? We think it can in spite of these curves, for two reasons. The first one is experimental, and we will presently turn to it. The second one is mathematical. We can show that a strictly two-phase model is able to generate the experimental results of Figure 2. This model will not be elaborated here. One obvious assumption it makes is that the phase of semantic activation and convergence is fairly short. Let us now turn to the further experimental evidence.

5. Are semantic alternatives phonologically activated? The problem resembles the problem of overlapping phases in the cohort model. There it was found that a phonologically activated item was subject to semantic context effects. This means that a word's phonological activation produces its semantic activation at a very early stage. In addition, the alternatives in the cohort also showed semantic activation. In other words, there was semantic unpacking of alternative items before the moment of convergence on a single candidate was reached.

We can now raise the corresponding issue for production, namely whether the semantic alternatives to our target item become phonologically activated. Is it, for example, the case that, when the target word is *table*, the word *chair* becomes momentarily phonologically activated in the speaker's mind? If this is the case, the situation is quite similar to the one in speech understanding: entering of the second access phase for all candidates *before* there is convergence on a single item.

How could this be tested? We used the same experimental procedure, though with different test probes. To investigate whether the semantic alternative *chair* became phonologically activated, we used a test probe like *char* (this is just an example; the experiment was run in Dutch). So, the subject would see a table to be named, but unexpectedly received the acoustic test probe *char*, right after the picture (there was only a short SOA condition in this case). We ran two versions of this experiment, and the results were crystal clear: there was no effect whatsoever. Probes that were phonologically related to semantic alternatives behaved exactly like the unrelated control items (such as *dog*).

Hence, the situation of word production differs essentially from that of word recognition. The alternatives to the target item do not enter the second

phase of access. In conclusion, if semantic alternatives to the target word are coactivated at the early stage of access, they are not phonologically unpacked to any measurable degree.

6. The process of phonological unpacking. It is well known from speech error research that the phonological form of a word is not retrieved in one swoop. The phonological encoding of a word proceeds rather through several steps (see Levelt 1989). Antje Meyer in our institute is at present completing an experimental project (Meyer 1988) on the course of such phonological encoding. The experiments were designed to investigate the temporal order in which different parts of a word form are activated, and to determine what kinds of phonological units (such as syllables, syllable constituents, and segments) are involved.

Meyer designed an ingenious technique for studying these matters. It is called 'implicit priming', and is exemplified in Table 1.

Table 1. Meyer's implicit priming technique. An example.

Homogeneous sets:	Cue word	Target word
Set 1	<i>touw</i> 'rope'	<i>kabel</i> 'cable'
	<i>poes</i> 'pussy'	<i>kater</i> 'cat'
	<i>peddel</i> 'paddle'	<i>kano</i> 'canoe'
Set 2	<i>bridge</i>	<i>poker</i>
	<i>stand</i>	<i>pose</i>
	<i>contract</i>	<i>polis</i>
Set 3	<i>cola</i>	<i>sinas</i>
	<i>graan</i>	<i>silo</i>
	<i>ring</i>	<i>sieraad</i>
Heterogeneous sets:		
Set 1	<i>touw</i>	<i>kabel</i>
	<i>bridge</i>	<i>poker</i>
	<i>cola</i>	<i>sinas</i>
.	.	.
.	.	.
.	.	.

A subject was presented with a small set of word pairs, for instance the homogeneous set 1 in Table 1, and had to learn these pairs. In some experiments there were sets of five word pairs, in others--like the one in the table--sets contained three word pairs. Each pair consisted of a cue word and a target word. These two always entertained an obvious semantic relation, which made it very easy to learn the pair. During the experiment the subject was given a cue word from the set and had to reproduce the corresponding

target word. The latency of this verbal response was measured. The cue words from a set were repeatedly presented to the subject, but in random order. After one set had been worked through, the subject was introduced to the next set, for instance set 2 of Table 1. The same testing procedure was followed for that set. And so on, until all sets had been run.

A homogeneous set is one in which the target words share some phonological property. In the example they share the first syllable. In set 1 this syllable is *ka*. In set 2 it is *po*; in set 3 it is *si*. But there were also heterogeneous sets. In a heterogeneous set the target words do not share the same property. One such set is given in the table (heterogeneous set 1). The target words have different first syllables. But notice that the pairs in this set are taken from the homogeneous sets. In other words, each target word appears both in a homogeneous set and in a heterogeneous set. Would the naming latency be different under these two conditions?

In the homogeneous condition the subject knows part of the target word before the cue word appears. That part may need no (or less) phonological encoding when the cue appears, and the response can be relatively speedy. Such preparation is impossible in the heterogeneous condition. If the two conditions show a difference in response latency, we can conclude that the implicit cue (e.g., the first syllable, as in the Table 1 example) has worked as a prime; the speaker could use it to encode part of the response. It turns out, for instance, that when the cue is the first syllable of the target word, as in Table 1, there is a strong priming effect. The speaker can apparently use that implicit first syllable cue to prepare his response.

In her experimental program, Meyer tested the effects of a wide range of implicit cues. Among them were first syllable, second syllable, onset of first syllable, rime of first syllable (both in one- and two-syllable words), first syllable plus onset of second syllable, first plus second syllable (in three-syllable words), and so on. It is, of course, impossible to review all these effects here. The reader is kindly referred to Meyer (1988), which is available on request. But it is possible to formulate two major principles which emanated from this work.

The first one can be called the 'word onset principle': for a cue to be effective, it must begin at word onset. Effective cues are, for instance, the onset consonant of a word's first syllable, the whole first syllable, the first syllable plus the onset of the second syllable, the first two syllables, and so on. Totally ineffective cues are, on the other hand, the rime of the first syllable (even in monosyllabic words!), the onset of the second syllable, the whole second syllable, etc. This finding shows that a speaker cannot encode part of the word before the preceding part has also been prepared. Phonological encoding is a strictly left-to-right affair. This brings us to the second principle.

The second principle can be called the 'additional constituent principle'. It says that the effectiveness of a cue increases with its size, but only if the additional parts are syllable constituents. Take, for example, the word *moraal* (*moral*). Here, word-initial *m* is an effective cue. It is a syllable onset constituent. For the same word, *mo* is a more effective prime. The additional part *o* is a syllable constituent, a rime. The cue *mor* is an even better prime. The additional part *r* is a syllable onset. But if the additional part is not a syllable constituent, there is no increase of effectiveness. Take the word

morgen (*morning*). Word-initial *m* is a good prime, but *mo* is *not* a more effective cue. The additional part *o* is not a syllable constituent in Dutch (Van der Hulst 1985), but part of an obligatorily branching constituent *or*. Only if that whole constituent is added, i.e. to make the cue *mor*, a more effective cue than *m* alone is obtained. This latter result is in excellent agreement with evidence from speech errors provided by Stemberger (1983), and with experimental results obtained by Treiman (1984).

These two principles support the following notion of phonological encoding: the speaker accesses the phonological structure of a word in serial order. This information is used to create a string of pronounceable syllables, each to be composed of syllable constituents. The speaker assembles such constituents, one after another, from the segmental materials that are being unpacked. Each assembled syllable corresponds to a phonetic or articulatory program, and the eventual phonetic plan of a word is a string of such syllable programs (see also Levelt 1989).

This strict serial encoding of a word's phonetic program is reminiscent of the cohort model where, as we saw, the first, phonological phase likewise proceeds in a strictly serial fashion from word onset. It is, however, by no means ascertained that syllable constituents play a similar role as units of speech perception.

References

- Brown, R., and D. McNeill. 1966. The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5. 325-37.
- Butterworth, B. forthcoming. Lexical access in speech production. In: *Lexical representation and process*. Edited by W. Marslen-Wilson. Cambridge, Mass.: MIT Press.
- Fromkin, V. 1971. The non-anomalous nature of anomalous utterances. *Language* 47. 27-52.
- Levelt, W.J.M. 1989. *Speaking: From intention to articulation*. Cambridge, Mass.: MIT Press.
- Marslen-Wilson, W. 1987. Functional parallelism in spoken word recognition. *Cognition* 8. 71-102.
- Marslen-Wilson, W., and L. Tyler. 1980. The temporal structure of spoken language understanding. *Cognition* 8. 1-71.
- Marslen-Wilson, W., and L. Tyler. 1981. Central processes in speech understanding. *Philosophical Transactions of the Royal Society London B259*. 317-32.
- Meyer, A. 1988. *Phonological encoding in language production: a priming study*. Unpublished doctoral dissertation. Nijmegen University, Nijmegen.
- Stemberger, J.P. 1983. The nature of /r/ and /l/ in English: Evidence from speech errors. *Journal of Phonetics* 11. 139-47.
- Treiman, R. 1984. On the status of final consonant clusters in English syllables. *Journal of Verbal Learning and Verbal Behavior* 23. 343-56.
- Van der Hulst, H. 1985. Ambisyllabicity in Dutch. In: *Linguistics in the Netherlands 1985*. Edited by H. Bennis and F. Beukema. Dordrecht and Providence: Foris. 57-66.
- Zwitserslood, C.M.E. 1985. Activation of word candidates during spoken-word recognition. Paper presented at the 26th Annual Meeting of the Psychonomic Society, Boston.
- Zwitserslood, C.M.E. 1989. *Words and Sentences: The effects of sentential-semantic context on spoken-word processing*.