Morphology in Speech Comprehension

Cover illustration: Clemens Wilmsen, Kerstin Mauth, Inge Doehring
Printed and bound by Ponsen & Looijen bv, Wageningen

Morphology in Speech Comprehension

een wetenschappelijke proeve
op het gebied van de
Sociale Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit Nijmegen,
volgens besluit van het College van Decanen
in het openbaar te verdedigen
op donderdag, 17 januari 2002,
des namiddags om 1.30 uur precies

door
Kerstin Mauth
geboren op 12 juli 1971 te Köln, Duitsland.

Promotor: Prof. dr. A. Cutler
Co-promotor: Dr. J.M. McQueen (MPI)

Manuscriptcommissie: Prof. dr. R. Schreuder
Prof. dr. J. Sereno (University of Kansas)
Prof. dr. P. Zwitserlood (Westfälische Wilhelms-Universität Münster)

The research reported in this thesis was supported by a grant from the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

Acknowledgments

It is amazing that during my PhD time I could hardly imagine that this book would be finished eventually. It is even more amazing how natural it feels now that it lies in front of me. Of course, as Anne Cutler once said “Writing a thesis is **not** one of the most natural things a human being could do ...”! But without her enthusiasm it would have been even more unnatural! I would like to thank my promoter Anne Cutler for the sparkles in her eyes when she talks about exciting results and for holding the comprehension group together in such a great way. In my struggle against the perfidies of a psycholinguists’ life (such as the invincible *Mauth*-effect) I always had an untiring and fearless fighter at my side: my co-promoter James McQueen. His help was invaluable and without him this book would not be what it is. His door was always open and he never got tired of twisting his mind over difficult data. And he managed to keep me going when doubts threatened to win over my optimism. Thank you James!

When I started my PhD almost four years ago I not only had to learn the mysteries of UNIX but I also had to acquire a new language and - even worse - set up experiments in that language. Several people helped me to manage these challenges: Dirk Janssen taught me all the little tricks and twists in AWK that make your life as an experimenter so much easier. Also, together with Femke van der Meulen, he finally succeeded in making me a considerably good speaker of Dutch. My thanks to Femke for patiently recording all the strange nonsense words that I needed to torture my subjects. And to the two of you: thanks for being such good friends throughout these years! I also owe thanks to Harald Baayen who ploughed through CELEX for several hours to get me started on my first project. Together with Petra van Alphen I created the materials for my categorization experiments and without her Dutch expertise, subjects would have been even more irritated by these strange sentences. Petra became a really good friend and I would like to thank her for special hours at the *Spaanse Dans* and good laughs and talks in the past two years! Tau van Dijk, Maarten Jansonius and Mirjam Broersma deserve special thanks for running the last word-spotting experiment while I was spending my half year *abroad* (Leipzig!).

Working at the MPI is much more than science and research. For me the relaxed atmosphere and the kind attitude of colleagues was always very special and the good spirit among the PhDs is really remarkable. Many thanks to the volleyballers, my PhD colleagues and my three roommates Andrea Weber, Nicole Cooper and Petra van Alphen who all contributed to a great time at the MPI. During my PhD period I had the opportunity to visit the Max Planck Institute of Cognitive Neuroscience in Leipzig for half a year and I would like to thank Angela Friederici for providing this opportunity. Anja Hahne was my EEG-ERP teacher and I would like to thank her for fruitful discussions and for giving me insights into this interesting research area. Monika Lück und Britta Stolterfoht were my salad-mates in Leipzig and together we endured pitying views from devoted canteen-eaters. Thanks to you for your warm welcome in Leipzig!

Special thanks to my paranimfen Joana and Petra who agreed on sharing the exciting hours before and during the day of my defense. We will have a lot of fun! Uli, I would like to thank for all his music and his friendship and Martin deserves thanks and respect for willingly sharing his house and his family with me.

I want to thank my parents for their support throughout all the years even though they sometimes wondered what this was all about!

And of course Clemens for his unbelievable trust in me and for being the way he is!

Unser Kopf ist rund, damit das Denken die Richtung wechseln kann.
Francis Picabia

Contents

Introduction

Chapter 1

Understanding spoken language means mapping arbitrary sounds onto meanings. In other words, the sensory input has to be matched against information that is stored in our memory. The part of our brain which contains all our stored lexical knowledge is referred to as the *mental lexicon*. Once we have accessed a certain word we can retrieve from the lexicon all the accompanying information, such as semantic, syntactic or pragmatic characteristics, which is linked to that specific word. On its way from our ears to the mental lexicon, the sensory input undergoes a transformation in the sense that once the signal has been filtered from surrounding noise it takes on a more abstract form. A common assumption among psycholinguists is that some sort of mediating representations are needed for the mapping process. At an early level of speech processing these mediating units code information about the signal itself: these are called the perceptual units. While psycholinguists seem to agree widely on the necessity of such perceptual units (but see also Gaskell & Marslen-Wilson, 1997 or Goldinger, 1998), the nature of these units is still a debated topic. Perceptual units are usually distinguished from lexical representations that are assumed to exist at a later level of processing and that code information about larger units such as words. Usually the flow of information through the system is described in terms of *activation* of these units.

The inherent features of the acoustic speech signal impose important constraints on this mapping process. Speech is - unlike written language - a temporal signal in which information unfolds sequentially over time. Moreover, much like handwritten scripts, speech is subject to a great deal of variability due to background noise, dialect, speaking rate, coarticulation etcetera. And speech is continuous. Thus, although we subjectively apprehend discrete words, boundaries between words are not reliably marked in the speech signal. This is not the case in written language, where white spaces between words unambiguously mark the offsets and onsets of words.

The goal of psycholinguistics is the unraveling of the cognitive processes that are responsible for the smooth and effective conversion of sound to meaning while taking these features of speech into account. Although those processes that lead to word recognition are commonly referred to as ’prelexical processes’, a clear differentiation between early - prelexical - processes and later - lexical - processes is difficult to achieve. For example, one of the first problems a listener is confronted with is the so-called segmentation problem, which is caused by the lack of unambiguous cues to word boundaries in the signal. Nevertheless, the location of those boundaries is essential for the recognition of words, since these are the units that are stored in our mental lexicon and that are the basic units which make up sentences. Thus, the segmentation of the incoming speech stream into smaller units is a prerequisite for word recognition and has to occur *prior to* word recognition. There is a large body of empirical evidence that segmentation is in part the result of parallel lexical activation and competition among lexical elements that are compatible with the acoustic input. At the same time, it has been demonstrated that prelexical segmentation cues such as phonotactic constraints or metrical structure are used by listeners to correctly locate word-boundaries (Cutler & Norris, 1988; McQueen, 1998). That means that both lexical and prelexical processes are involved in the segmentation of speech.

Another example of the interdependence of prelexical and lexical processes is the way in which decisions about phonemes can be determined by prelexical - phonemic - information as well as lexical, semantic and syntactic information. If listeners are confronted with an ambiguous sound they not only use phonemic information but also use contextual information to come to a decision about that sound. This thesis will concentrate on these two topics in spoken-word recognition. The segmentation problem is addressed in Chapters 2 and 3, while phonemic decision-making is investigated in Chapters 4, 5, and 6.

Models of spoken word recognition

Within the discipline of speech comprehension different models have been developed which have tried to incorporate the mechanisms and strategies that listeners use to convert sound into meaning. In the following section the most influential models will be briefly introduced. Although the architectures of these models differ to varying degrees, they all agree on three basic assumptions: there are mediating form representations of some kind, there is multiple parallel activation of lexical elements, and there is competition among these lexical representations for recognition.

COHORT

The COHORT model (Marslen-Wilson & Welsh, 1978) was one of the first models that was explicitly designed to account for how the inherent features of spoken language are processed. Lexical knowledge in the COHORT model is represented by memory elements in the form of processing units that actively respond to sensory information of the signal. Word recognition is achieved in the following way: on the basis of the first sound of the stimulus input, all words that match that input are activated to form the initial cohort. Thus, the perception of the sound [m] activates - among others - the words *method, medal* and *miracle*. This selection of a first set of candidate words is purely data-driven in the sense that it is exclusively determined by the acoustic-phonetic properties of the input. A preselection of cohort members based on contextual information is not possible in this model. Once the initial cohort has been formed, all activated candidate words in that cohort monitor the subsequent sensory input and remove themselves from the cohort as soon as mismatching input information is encountered. The perception of the sound [ɛ], for example, will cause the word *miracle* to be dropped from the cohort while the words *method* and *medal* will remain in the cohort. Word recognition can occur as soon as the initial cohort has been reduced to one single member. The exclusion of candidate words from the cohort is also determined by contextual information since lexical representations can receive information about the semantic and syntactic appropriateness of the words they represent.

Although COHORT successfully incorporates two major features of the recognition process, namely multiple activation and competition of lexical candidates, there is one serious problem the model cannot solve. It does not provide a mechanism that allows words to get back into the cohort once they have been excluded from it or that allows words into the cohort when the beginning of the word has not been apprehended correctly. To cite the well-known example from Norris (1982, 1994), the word ’cigarette’ would never enter the cohort when pronounced ’shigarette’. This radical consequence of mismatching information on word recognition is not reflected in human language comprehension, which is flexible with respect to mismatching information and which can recover from mispronounciations like the one mentioned above. Furthermore, since COHORT puts strong emphasis on the onsets of words, the continuous nature of spoken language is problematic for this model: word boundaries - and therefore word-onsets - are not reliably marked in the input and therefore the listener might be uncertain about the beginning of a word in the continuous speech stream.

Recently, a revised version of the COHORT model, namely the *Distributed Model of Speech Perception* developed by Gaskell and Marslen-Wilson (1997), has taken a very different approach to spoken language processing from that in the original COHORT model. Lexical knowledge in this account is not locally represented by discrete units for each lexical entry, but rather by distributed representations using the same nodes for all lexical entries. The only means by which words can be discriminated is the pattern of activation and the amount of activation across the array of lexical representations at particular moments in time.

TRACE

TRACE (McClelland & Elman, 1986) was designed to overcome the drawbacks of cohort theory whilst keeping its merits. In contrast to COHORT, TRACE is an entirely interactive network that allows information to flow both bottom-up and top-down. It operates on three processing levels that consist of units representing features, phonemes, and words respectively. These units process information via bidirectional excitatory or inhibitory connections whereby each unit has to reach a certain level of activation before it can start spreading its activation to other units. The units of the network are viewed as hypotheses about the input that is currently being processed and are continuously constrained by each other. In order to account for the continuous nature of the speech signal, the entire network is copied over and over again so that all states of activation are available at each point in time where a word might start. Since onsets of words are not reliably marked, all phonemes in the input are potential onsets, and thus the input *tree* will not only activate words that start with [t] (as would be the case in COHORT) but also words that start with [r] and those that start with [iː].

The successful selection of one single word that matches the input best is achieved by inhibitory connections between units that represent words at the lexical level. Words that overlap in time inhibit each other, thus reflecting the competition process between words that are compatible with the input. The current state of activation of each unit at each of the three levels is determined by the inhibitory or excitatory information it receives from those units it is connected to. Thus, hypotheses that are active at the phoneme level in TRACE are not only determined by data-driven information but are also strongly influenced by lexical information that is fed back from the lexical level.

One major problem of TRACE is that it requires massive duplication of units and connection patterns. This severely constrains the size of the lexicon it can be applied to. Furthermore, the encoding units in TRACE are at the same time decision units from which phonological output is read out once information processing has been finished. Since the encoding units are continuously updated, there is no way in TRACE to keep a record of previous decisions. Once a decision has been made there is no way to alter that decision by subsequent information. The consequence is that TRACE is not able to identify mispronounciations since there are no continuing representations of earlier decisions. Finally, the notion of feedback has been questioned both on empirical and theoretical grounds. A detailed discussion of the feedback issue will be provided in Chapter 4.

The Neighborhood Activation Model

The Neighborhood Activation Model (NAM; Luce, Pisoni, & Goldinger, 1990) puts its emphasis on the influence of neighborhood activation and neighborhood density on word recognition. The neighborhood of a given lexical entry is simply the population of words that are phonetically similar to that lexical entry. The most important variables that determine the lexical access process are - besides word frequency - neighborhood density (i.e., the number of phonetically overlapping lexical entries) and neighborhood frequency (i.e., the summed frequency of phonetically similar entries). The NAM - like COHORT and TRACE - is an activation model and shares with COHORT the assumption that the initial activation of word-decision units stored in memory is purely data-driven since this activation is directly and exclusively based on the activation of acoustic-phonetic patterns. It assumes - much like COHORT and TRACE - that there is multiple activation of representations that are compatible with the sensory input. As more input is provided to the system, the activation levels are adjusted such that those patterns which still match the input are more strongly activated while the activation of similar patterns which diverge from the input is attenuated. Once a word-decision unit has been activated, it is modulated not only by bottom-up information provided by the acoustic-phonetic input but also by the overall activation state of the decision system as well as by word frequency. The output of the decision system (i.e., the selection of a single word) is computed by a so-called ’neighborhood probability rule’ which considers the following values, which are weighted relative to each other: the frequency of a stimulus word and its intelligibility, the neighborhood confusability (i.e., how many similar words are competing for recognition), and the frequency of the relevant neighborhood. Thus, word recognition is based on the computation of the likelihood of one lexical entry to be favored above its neighbors.

The model differs from COHORT in the sense that the computation of a recognition value is responsible for the selection of the intended word rather than the reduction of the cohort down to a single candidate. It also differs from TRACE in that the NAM does not rely on inhibitory connections between lexical representations within the lexical level to choose the best-fitting candidate word. PARSYN, a connectionist network version of the NAM, has recently been developed (Luce, Goldinger, Auer, & Vitevitch, 2000) in order to simulate the predictions of the NAM.

Shortlist

Like COHORT and the NAM, Shortlist (Norris, 1994; Norris, McQueen, Cutler, & Butterfield, 1997), is a data-driven model. It is more radical than COHORT and the NAM, though, since it assumes not only that the initial processing stage is data-driven, but also that information in the entire system can only flow bottom-up so that higher order information can not influence processes at earlier processing stages. The process of word recognition in Shortlist starts with an exhaustive search of the entire lexicon. This generates an initial list of potential word candidates (the *shortlist*) that begin at each phoneme in the input. Whether or not a candidate word enters the shortlist depends on the goodness-of-fit of that word with the input. The size of the shortlist is regulated such that once there are too many candidates in the shortlist, those with the least bottom-up support are eliminated in order to allow candidates with more bottom-up support to move up into the shortlist. The shortlist of potential candidates is then passed on to an interactive activation network where all candidate words compete with each other for recognition. This second stage works just like the lexical level in TRACE: words overlapping to a certain degree inhibit each other proportional to the number of phonemes they share. Since the number of competing elements is restricted and thus the number of connections that mediate activation is also restricted, Shortlist can handle a lexicon that is of a realistic size (i.e., including as many as 25,000 words).

Although Shortlist shares with TRACE the fact that the segmentation of continuous speech results directly from the competition process, some basic assumptions are fundamentally different between the models. First, while TRACE allows information to flow both bottom-up and top-down, Shortlist assumes strictly bottom-up processing. This issue has recently been addressed by Norris, McQueen, and Cutler (2000) who extended Shortlist with the Merge model. Merge models the way listeners make explicit phonemic decisions and shares the assumptions made by Shortlist including the assumption that there is no feedback from the lexicon to lower levels of processing. The Merge model and the feedback debate will be discussed in Chapter 4.

The second difference between Shortlist and TRACE is that the former makes use of explicit form-based representations of words while the latter does not have such representations. In TRACE, the form information of a particular word is only coded in the links between that word's lexical node and lower level phoneme nodes, which also encode the input. The advantage of these explicit representations in Shortlist is that the system - unlike TRACE - can not only detect mispronounciations but can also identify them since a direct comparison of the actual input can be matched against stored lexical representations.

Shortlist in its original version accounted for segmentation using competition alone: in the German sentence '*Sein Kaufrausch war nicht zu bremsen*' (*His buying addiction could not be stopped*) words like *ein (one), auf (on), Frau (woman), rauh (rough)* and *ich (I)* will temporarily be activated. But finally the appropriate words will win the competition because they provide a complete parse of the input while the other candidates would leave parts of the input unaccounted for. Recent research, however, has demonstrated that competition is not the only means by which the speech stream can successfully be segmented. These segmentation strategies will be described in more detail in Chapter 2. Shortlist has been adapted such that the competition process can now be modulated by various segmentation cues.

Morphological issues in spoken-word recognition

Morphological issues have played an important role in psycholinguistics for a long time. Linguistically speaking, morphemes are those combinatorial units that make up the words that we know and they are thus the basic units that form our lexical knowledge. Thus, no matter how morphologically simple or complex a given language might be, the investigation of morphologically complex words can give us insight into the structure of the mental lexicon. In the lexical access process, for example, morphemes might play an important role in the process of sound-to-meaning conversion. It might be possible that morphemes rather than whole words form the mediating representations between the sensory input and the more abstract mental lexicon. One of the most hotly debated topics in the field of psycholinguistic research concerns exactly this issue: what is the nature of the representations used to access the lexicon, what is the structure of the linking process between access and more central representations in the lexicon and how are the underlying lexical entries in the mental lexicon organized? Is morphology as described by linguists directly reflected in the organization of the mental lexicon and the access process, or is morphology rather a linguistic tool which can be used to describe the structures of words but which has no role to play in the processing of speech? This thesis will be about this issue, specifically the role of morphology in spoken word recognition, focusing on the two topics mentioned earlier: the segmentation of the continuous speech stream and phonemic decision-making.

Morphological relationships between words fall into two classes: while inflectional morphology codes, for example, number or tense information - as in *flowers* and *laughed* respectively - derivational morphology forms new words usually of a different word category such as the adjective *sunny* from the noun *sun*. Research about morphological issues in language processing has involved both derivational and inflectional morphology but most of that research has concentrated on the visual rather than the auditory domain. As already mentioned, the underlying processes that convert the auditory input into meaning are not necessarily the same processes that convert written input into meaning. Furthermore, there might be different processes underlying inflectional and derivational morphology. The current work is concerned only with the role inflectional morphology might play in auditory word recognition.

Access versus central representations

In the study of language comprehension (either spoken or written) one of the most important distinctions to be made is the one between *access* representations and *central* representations. The former are modality-specific and are understood as mediators between the perceptual input and the mental lexicon because they code form information (either phonological or orthographic). To link back to the previous section in which models of speech processing were discussed, these access representations are those units that are referred to as *lexical* representations in models like Shortlist, TRACE or COHORT. Access representations are thus located at what is called the *lexical level* in comprehension models. Central representations, on the other hand, are seen as form- and therefore modality-independent bundles of all the information (including semantic and syntactic features) relevant to a given lexical entry. These elements thus form the core *mental lexicon*.

While researchers agree on this distinction, their opinions are still divided on the following questions: What are the units of representation? Is the mental lexicon organized in terms of morphemes rather than in terms of words? That is, are morphological relationships represented independently of form- or semantic information? Do access representations reflect the structure of the mental lexicon, that is, are access representations structurally ”defined”in the same way as central entries? In order to provide answers to these core questions, researchers have concentrated on the investigation of the processing of morphologically complex words, because psycholinguistic insight into the processing of those forms allows one to develop a more detailed view of the organization of the mental lexicon (Marslen-Wilson, Komisarjevsky Tyler, Waksler, & Older, 1994).

Full listing versus full parsing

One of the most important issues concerning the processing of morphologically complex words is whether *decomposition* (i.e., morphological analysis) of complex words prior to lexical access is obligatory, optional or impossible. In other words: do we use morphological rules during processing whenever this is possible and thus store only that information which is not covered by these rules (e.g., the decomposition of regularly inflected forms vs. the storage of irregularly inflected forms; Clahsen, 1996)? The same question can be asked about the central representations: are morphological relationships between stored lexical entries represented in the mental lexicon? Note that it is possible that the lexical access process might be mediated via morphological decompositional routines while at the same time the central representations might not reflect this morphological structure. The opposite might also hold: while full-form representations at the lexical level might serve as mediators between sensory input and the mental lexicon, the lexicon itself could be structured along morphological principles.

Two extreme positions have been contrasted in this line of discussion: *full parsing* versus *full listing*. The former implies absolute transparency of all morphologically complex words, stating that each morpheme needs to have its own access representation that maps form onto meaning. The latter assumes the opaqueness of complex forms, implying that each complex form has its own access representation regardless of its morphological structure. Similar assumptions hold for the central lexical representations. Seen from a full parsing perspective, the idea is that the lexicon is organized along morphological lines and that there are connections between those units that can be combined with each other (Jarvella & Meijers, 1983), while a full listing account assumes a list of whole forms in the lexicon which does not reflect their morphological complexity (Butterworth, 1983).

Inflectional morphology and lexical access

A review of the recent literature shows that neither of these two extreme positions can be upheld, at least not in their strongest versions. The full parsing account as put forward by Taft and Forster (1975), for example, has been challenged by surface frequency effects observed for inflected word forms like, for example, the Dutch plural form ’wolken’ (*clouds*; Baayen, Dijkstra, & Schreuder, 1997; Baayen, McQueen, Dijkstra, & Schreuder, 2001). The surface frequency of a word is the sum of the occurrences of a given word-form (e.g., of the form *wolken*). In an auditory lexical decision task, Baayen et al. (2001) presented two groups of noun plurals (such as *wolken*) that were matched on their combined stem frequency (i.e., the combined frequency of all words that contain the words' stems) but that varied with respect to their plural frequencies. Thus, one group of items was singular-dominant (i.e., the singular form was more frequent; e.g., the singular *soep* '*soup*' is more frequent than the plural *soepen* '*soups*') while the other group was plural-dominant (e.g., *wolken* is more frequent than *wolk*). The authors observed faster reaction times (RTs) in auditory lexical decision on plurals for plural-dominant forms (i.e., *wolken*) as compared to singular-dominant forms (i.e., *soepen*). When, however, the singular forms were presented (i.e., *wolk* and *soep*, respectively) no RT difference was observed suggesting that the recognition of the singular forms was influenced by the combined stem frequency while the recognition of the plurals was influenced by the surface form frequency.

Such a surface-frequency effect - or missing base-form frequency effect - for regular plurals cannot be explained within a full-parsing model since it requires the decomposed representation of morphologically complex forms. Such a model would predict that both plurals and singulars are accessed via their stems and therefore that only combined stem frequency should modulate word recognition, not surface frequency.

Taft (2001) recently argued, however, that surface frequency effects are not necessarily evidence for full-form representations and against decomposed representations. He asked whether the surface frequency effect observed for plural dominant nouns, for example, might not in fact reflect the ease of processing of a given stem-morpheme combination. Since the plural form *wolken* is more frequent than the plural form *soepen*, the likelihood of the combination of the stem *wolk* and the plural affix *-en* is higher than the combination of the stem *soep* with the affix *-en*. According to such an account, both forms would be accessed via their stems and the advantage of the higher frequent form would only be effective at a later stage where the evaluation of the likelihood of the (re)combination with the plural affix would be easier for a more frequent plural form. Note that this assumption implies that frequency is a characteristic that is stored at the central rather than at the access level.

The results reported by Baayen et al. (1997, 2001) were also taken to challenge full-listing accounts (Butterworth, 1983), although the surface frequency effect for inflected forms fits such an account perfectly. Remember, however, that the words used in that study were matched on their combined stem frequency but that their surface frequencies were varied. Thus, one group of words was singular-dominant while the other group was plural-dominant (i.e., the surface frequency of the singular form *soep* was more frequent than the singular form *wolk*). Baayen et al. did not find surface frequency effects for these singular forms. They therefore concluded that the singular forms were accessed via stem representations that are sensitive to combined frequencies of singulars and plurals while the plural forms were accessed via full-form representations that are sensitive only to plural frequency. A full-listing hypothesis cannot explain this lack of a surface frequency effect for singular forms.

Therefore, models ranking between the extreme positions of *full parsing* and *full listing*, including those which have been called dual-route models, are much better suited to explain results like those reported above. These models assume only partial decomposition depending on the morphological structure and features of morphologically complex words. The idea in a dual-route model is that listeners have two different processing strategies at hand, and that the ”choice”of the better (i.e., the faster) one depends on the morphological structure of the form to be processed. Schreuder and Baayen (1995), for example, developed a Race model in which two routes - a direct mapping route and a parsing route - work in a fully parallel fashion. The model consists of a spreading activation network with three representational levels where modality-specific access representations are distinguished from integration nodes and modality-independent semantic/syntactic representations.

Processing via the parsing route involves three stages and is therefore in general slower than the direct mapping route. The direct mapping route directly maps a full-form representation to the corresponding integration node and then via activation directly to the syntactic/semantic representation. Both routes are available and act simultaneously during information processing: hence the term Race model. Fully transparent forms might be processed by either of these routes. Parsing will only win the race, however, when it is fast. But since parsing involves more processing steps than the direct mapping route it tends to be more time consuming than direct mapping. This encourages the development of full-form representations which can be recognized via the direct route. The model therefore allows for the full-form representation of morphologically complex yet transparent forms. In other words, if a given regular wordform is frequent enough it might be more economical for the system to store that word’s form and to access it via a look-up process rather than decompose it before lexical access every time it is encountered.

Another dual-route model has been put forward by Clahsen (1999) and Clahsen, Eisenbeiss, and Sonnenstuhl-Henning (1997), who suggest a divergence of two access routes along more linguistically motivated lines. According to their hypothesis, the parser distinguishes between regular and irregular wordforms and makes use of that information in the choice of the appropriate access route. The logic is that the parser can make use of rules to decompose and access regularly inflected forms while these rules cannot be applied to irregularly inflected forms that instead need to be stored as whole wordforms. In a visual lexical decision experiment, Clahsen et al. (1997) reported word form frequency effects for German irregular participles (*gelaufen*, ’ran’) as compared to regular participles (*gelacht*, ’laughed’) whose surface frequency was matched to that of the irregular participles. The authors took this result as evidence for storage of irregular forms because only if these forms are stored can the influence of their surface frequency on recognition performance be explained. Because no such effect was found for the regular forms, the authors concluded that these forms have no full form representation in the lexicon and thus have to be accessed via a decomposition process.

Note that the dual-route model of Clahsen et al. (1997) cannot explain the surface frequency effects for fully regular plurals in Dutch observed by Baayen et al. (1997, 2001) since, according to the model, regular plurals should be processed via the decompositional route and should therefore not be sensitive to surface frequency effects. However, these two dual-route models share the assumption that some inflected forms might have independent full-form representations. They moreover agree in the observation that morphological information is represented to some extent in the mental lexicon.

Further support for the representation of inflectional morphology in the mental lexicon comes from an investigation in Italian by Caramazza, Laudanna, and Romani (1988). These authors reported systematic influences of morphology on RTs and error performance in three lexical decision experiments. The experiment measured how easily subjects could reject nonwords that were decomposable into morphemes to various degrees. The authors found that subjects were fastest in rejecting nonwords that were not decomposable into smaller units (like the Italian nonsense string *canzovi* that contains neither a stem nor a verbal affix). It was harder for subjects to reject nonwords that consisted of at least one legal morpheme (like the nonword *cantovi* which contains the legal stem *cant-* (to sing) but ends with an illegal Italian suffix *ovi*). Hardest to reject were nonwords that were constructed illegally out of two existing morphemes (like the incorrect form *cantevi* that can be decomposed into the stem *cant-* and the suffix *-evi*). If listeners had not employed some kind of decomposition strategy there should have been no observable RT or error rate differences between the three types of nonwords.

Caramazza et al. (1988) took these results as support for the Augmented Addressed Morphology Model (AAM), which assumes that lexical access for known words might take place through a direct whole-word mapping procedure while novel words might be processed by addressing representations of morphemes.

Although the model was primarily designed for visual word recognition, similar assumptions were considered possible for the auditory domain. It is much like the Race model (Schreuder & Baayen, 1995), that is, it is an activation-based model which assumes the parallel activation of both full-form representations (e.g., *walked*) and the respective morpheme constituents (i.e, *walk-* and *-ed*) as well as (orthographically) similar forms (e.g., *walks, walking, talking, winked*). Recognition can occur as soon as a certain activation threshold has been exceeded. The authors assumed that full-form representations (when available to the reader due to earlier exposure) should reach the critical threshold faster than decomposed representations. This full-form representation then makes contact with the decomposed representation of the morphologically complex form (i.e., the full form *walked* mediates information to the two morphemes *walk*$_V$ $+$ *-ed*). Since the activation of a full-form in the AAM depends on the familiarity of such a form, novel words can only be processed via a decompositional route since no full-form representation is available for them. Processing takes longer in that case because computation rather than direct mapping is required. One major criticism of the AAM was that it was in part based on observations made about the processing of nonwords. This weakens the generalizations that can be made about the processing of existing words.

Derivational morphology in spoken-word recognition

Although the current study focusses exclusively on the processing of inflectional morphology, work on spoken language with derived forms also has important implications for models of morphological processing. Marslen-Wilson et al. (1994), for example, used the cross-modal priming task in order to investigate the relative contribution to the representation of derived word forms of phonological and semantic transparency, on the one hand, and morphological relatedness, on the other hand. In cross-modal priming, the prime is presented auditorily and is then followed by a visually presented target on which the subjects are asked to perform a lexical decision. A priming effect is observed when there is a relationship of one kind or the other between prime and target. For example, the presentation of the prime *rose* might speed up lexical decision on the target *flower* since the words are semantically related. Because in this task prime and target are presented in different sensory modalities, effects observed with this task are supposed to tap into more central processing stages since there are no shared form representations that might mediate priming effects. Marslen-Wilson et al. (1994) observed different priming effects depending on the derivational relationships between primes and targets. They found priming effects for suffixed words and their stems irrespective of which was presented as prime or target. Thus, the suffixed form *friendly* primed its stem *friend* and was also primed by it. The same was true for derivational words that were prefixed: *unhappy* primed *happy* and was primed by it. While prefixed derivational forms that shared the same stem also primed each other (e.g., *unfasten* primed *refasten* and vice versa), this was not true for suffixed derived forms. There was thus no priming effect between the two derived forms *confessor* and *confession*.

These results were interpreted as evidence for a morphologically-structured mental lexicon in which stems have links to those morphemes they can combine with. Due to the decompositional structure of the mental lexicon, priming effects occur because access to morphologically complex forms is always mediated via the stem. Thus, repeated activation of the stem leads to the observed priming effects. But if so, why was there no priming effect for suffixed words derived from the same stem? Marslen-Wilson et al. (1994) proposed that this was due to inhibitory links between suffixes. Since all possible affixes should be activated by the stem they are linked to, successful selection of the appropriate suffix can only take place when all other competing suffixes are suppressed via inhibitory activation. This sort of inhibition mechanism is not required for prefixes because - due to their different onsets - they should not be activated simultaneously by the same input. As McQueen and Cutler (1998) argue, the assumption of inhibitory links between derivational suffixes might also be applied to inflectional morphemes. A shared-entry model with between-entry connections that varied in strength - such as the one proposed by Marslen-Wilson et al. (1994) - could account for any priming asymmetries between morphologically related words.

Recently, the claim about inhibitory links between derivational suffixes in the mental lexicon has been challenged by a study in German run by Zwitserlood (2001). She reported data from both unimodal and cross-modal priming studies that showed not only reliable priming effects between two derived forms that shared the same stem (e.g., *Achtung*, ’attention’, and *achtsam*, ’attentive’) but also between two inflected forms (e.g., *glauben*, ’believe’, and *glaubte*, ’believed’) and between inflected and derived forms (e.g., *hungrig*, ’hungry’, and *hungerte*, ’she/he hungered’).

Note that the priming effects that Marslen-Wilson et al. (1994) observed for prefixed words that shared the same stem can be explained, at least in principle, by both continuous and discontinuous models of auditory word recognition. A discontinuous (i.e., decompositional) model (Taft & Forster, 1975) would account for these priming effects by assuming that the prefix has to be stripped off the stem so that word recognition via that stem can occur. Such a model is discontinuous in the sense that it assumes that information which could be available to the system early (e.g., the prefix *re-* in *refasten*) is ’delayed’ until the stem has been accessed and hence the combined meaning of prefix and stem can be recovered. A continuous model, on the other hand, assumes that the auditory input is processed in a left-to-right manner so that incoming information is mapped continuously onto representations in the mental lexicon (e.g., Shortlist, COHORT or TRACE). Such a model can explain the priming data for prefixed words only if it assumes that some morphological information is stored at the lexical level: once a prefix is activated, this activation spreads to the linked stem(s) and hence priming can occur to the same stem preceded by another prefix.

Schriefers, Zwitserlood, and Roelofs (1991) explicitly compared continuous with decompositional models by investigating Dutch prefixed words and pseudo-prefixed words. While prefixed words contain a free stem (e.g., *heading* in *subheading*) pseudo-prefixed words do not (e.g., **ject* in *subject*). In a first experiment, Schriefers et al. (1991) compared phoneme monitoring latencies in three different conditions: a first condition in which the phoneme that the subjects had to monitor for was contained in a free stem (e.g., [n] in *staan*, ’to stand’); a second condition in which the phoneme had to be detected in a derived relative of the stem (e.g., [n] in *opstaan*, ’to get up’); and a third condition in which the phoneme had to be detected in another derived form of the stem (e.g., [n] in *toestaan*, ’to allow’) which differed from the other two conditions in that its uniqueness point (UP) was at least one phoneme before those of the other two conditions. While [n] is the UP of both *staan* and *opstaan* - i.e., up to that point there are still other competitors that are also compatible with the input - the UP of the derived form *toestaan* is the [a:]. The authors based their experiment on earlier findings with phoneme monitoring that showed that the detection of a previously specified phoneme is faster when that phoneme occurs after the UP then when it occurs before that point (Frauenfelder, Seguí, & Dijkstra, 1990). If prefixes are stripped off their stems prior to lexical access (as is predicted by the decompositional account), phoneme monitoring should be equally fast in all three conditions since all three words should be accessed via their stems. However, if prefixed words are processed in a strictly left-to-right fashion, monitoring latencies should be faster in the word with the earlier UP (i.e., *toestaan*).

The authors observed faster monitoring latencies for both prefixed derived forms (i.e., *opstaan* and *toestaan*) as compared to their stem. While this finding was clear evidence against an affix-stripping approach, it was less easily interpretable with respect to continuous approaches because one important factor - namely, that of size of the competitor set - was not kept constant between the conditions. Since the words in the three conditions all had different onsets, they activated different sets of competitors. Therefore, the overall advantage of prefixed words over stems could have been due to a larger set of competitors for the stems. In a second experiment, Schriefers et al. (1991) controlled for that factor and found the same result as in the first experiment: prefixed words elicited faster RTs than the corresponding stems. The authors therefore concluded that the UP of prefixed words was not a good predictor of the recognition point of those words.

Schriefers et al. (1991) also tested whether the same finding would be obtained for pseudo-prefixed (i.e., monomorphemic) words such as *subject*. A decompositional model would predict that listeners would attempt to recognize the pseudo-prefixed words via the bound stem and that a reanalysis of the input would become necessary at the point of the bound stem when no other free stem would be compatible with the auditory input. For example, only at the phoneme [k] of the bound stem **zwijken* in *bezwijken* (to give way) can the listener start a reanalysis, since up to that point other free stems would still compete for recognition (e.g., *zwijn*, 'swine', or *zwijgen*, 'to be silent') . Schriefers et al. (1991) called that point in the pseudo-prefixed words the *pseudo-stem point*. The UP of the whole word *bezwijken* is already at the [ɑɪ], thus one phoneme before the pseudo-stem point. Phoneme monitoring latencies were compared for pseudo-prefixed words and prefixed words. These were matched such that the UP of the prefixed words was the same phoneme as the pseudo-stem point of pseudo-prefixed words (e.g., the [k] in the prefixed word *bestrijken*, 'to strike', and in the pseudo-prefixed word *bezwijken*). Listeners had to monitor for these sounds. Continuous models would predict that RTs should be faster to pseudo-prefixed words than to prefixed words since the relevant phoneme always followed the UP in these cases. The decompositional account would predict the opposite since the reanalysis required in pseudo-prefixed words should lead to an increase in RTs. The results were clearly in favor of continuous accounts: phoneme latencies were faster for pseudo-prefixed words than for prefixed words.

On the basis of these results, Schriefers et al. (1991) concluded that listeners did not make use of a decompositional strategy before they accessed the words but instead that auditory word recognition proceeds in a left-to-right fashion. Nevertheless, they argued that the identification of morphologically complex words was not purely based on full-form access. They argued that the mental lexicon must also contain information about the constituent parts of complex words in order to accommodate the observed results (i.e., the faster responses to [n] in *opstaan* than in *staan*).

Recently, Wurm and Ross (2001) introduced a construct called the *conditioned root uniqueness point* (CRUP) and proposed that some of the results obtained by Schriefers et al. (1991) might be influenced by this factor. The CRUP of a prefixed word is that point in the word when - once the prefix has been stripped off - no other free stems compete with the free stem of the prefixed word in question. For example, the prefixed word *relive* competes with other words like *religion* or *relinquish* and its full-form UP is at the [v]. However, when the prefix is stripped off, no other free stem that starts with the phoneme [l] can combine with the prefix *re*-. Thus, the CRUP of the complex word *relive* is the [l] of *live*, which precedes the full-form UP [v]. Wurm and Ross (2001) used different tasks (gating, lexical decision, and naming) in order to test how far auditory word recognition might be affected by CRUPs. They found an overall RT advantage of CRUP words over non-CRUP words (i.e., those words in which UP and CRUP fell on the same phoneme). They therefore proposed a dual-route model in which words can be analysed via full-form representations or via their constituent parts when possible. The decompositional route differs from that in other models in that - after an affix has been stripped off its stem - only those lexical entries that are compatible with the acoustic signal, and, moreover, are free stems that are compatible with the given prefix, will be considered further. The set of competitors is thus reduced from the entire lexicon to only a small number of items that fulfill certain morphological constraints. A re-analysis of the Schriefers et al. (1991) data that was sensitive to CRUPs might explain those data within this new framework.

To summarize, there is clear evidence that the processing of spoken language is sensitive to morphological information. However, the processing of morphologically complex words seems to be influenced by more factors than mere linguistic specifications. Among these factors (which have not all been discussed here) are surface- and combined stem frequencies (Baayen et al., 1997), family size (de Jong, Schreuder, & Baayen, 2000), order of occurrence of affixes and stems (Marslen-Wilson et al., 1994), semantic transparency (Marslen-Wilson et al., 1994), and the regularity of the morphological process (Clahsen et al., 1997). It is also important to note that the balance of storage and computation of complex forms might also depend on the morphological complexity of a given language. While Dutch is morphologically rather simple, Finnish, for example, has a very rich inflectional morphology allowing for hundreds of inflectional variants of one single word (Karlsson & Koskenniemi, 1985). The storage of all these variants seems uneconomical since this would lead to an enormous mental lexicon. And indeed, so far hardly any empirical evidence has been reported that supports lexical storage of inflected words in Finnish (Niemi, Laine, & Tuominen, 1994; Bertram, Laine, & Karvinen, 1999; Bertram, Schreuder, & Baayen, 2000). The relationship between models of morphological processing and models of spoken language comprehension is likely to be affected by the morphological richness of a given language.

To date, no model of spoken language processing (e.g., COHORT, TRACE, NAM, or Shortlist) has been adapted to take morphological information into account. The experiments described in this thesis contribute to our understanding of how morphological information can be integrated into these kinds of models. As will become clear in the following chapters, morphological issues can have interesting implications for the architecture of models of spoken language processing.

The current study

The present study examined the role morphology might play in early stages of speech processing. In that respect, two different aspects of speech processing were investigated. The first part of the thesis (Chapters 2 and 3) is concerned with whether morphology plays a role in speech segmentation, while the second part (Chapters 4, 5, and 6) investigated how and when morphological information is used in speech recognition as measured with the phonetic categorization task.

As already mentioned, continuous speech confronts the listener with the problem of segmentation: in order to map sound onto meaning, one of the first tasks of a listener during speech processing is to identify the words that make up the continuous utterance. Since there are no reliable cues to word boundaries in the auditory signal, this task is far from trivial. However, a considerable amount of evidence has been obtained in the last few years which suggests that there are many mechanisms which listeners use to solve the segmentation problem. A complete overview of this literature will be provided in Chapter 2. Among these mechanisms is the Possible Word Constraint (PWC; Norris et al., 1997) which states that listeners segment the speech input such that no impossible words are left over in the ongoing parse of the input. The stretch of speech which constitutes a *possible word* in a given language will also be discussed in more detail in Chapter 2. For many European languages, however, it holds that a possible word has to contain at least a vowel. This implies that a single consonant is not sufficient to form a possible word in Dutch, for example. However, inflectional morphemes often consist of single consonants and thus the question arises whether the PWC might be sensitive to that information. So far, the PWC has been proposed to be a purely phonological mechanism not sensitive to higher-order information. Experimental evidence as to whether the PWC is sensitive to morphological information or not also has interesting implications for theories of morphological processing. If one could demonstrate experimentally that morphological information influenced the processing of the PWC, this would imply that morphological structure was represented at a very early stage in speech processing. Experiments on the status of inflected morphemes in segmentation will be presented in Chapters 2 and 3.

The second part of the thesis will be concerned with potential influences of morphological information on phonemic decisions. The relative contributions of different sources of information (phonemic, lexical, or syntactic) on phonemic decisions are usually assumed to shed light on the architecture of models of spoken word recognition such as TRACE or Shortlist. The sources of information which listeners use at different points in time during language processing to identify the sounds that make up the acoustic signal will be described in Chapter 4. The experiments reported in Chapters 4, 5, and 6 all investigate whether morphological information has a role to play when listeners are explicitly asked to identify an ambiguous sound as one of two phonemes. Inflected verbs and uninflected nouns will be presented both in appropriate sentence contexts and in isolation in order to examine how morphological information is integrated during sentence processing.

The Possible Word Constraint and inflectional morphology

Chapter 2

Segmentation strategies across languages

In recent years researchers in the field of psycholinguistics have paid a great deal of attention to the issue of how listeners master the very complex task of segmenting the auditory speech input into recognizable units. It is well known that the speech stream provides few cues for the boundaries of those units that we subjectively apprehend as separate words. The impression of hearing discrete words one after the other must therefore be the consequence of the recognition process.

Some word boundaries, admittedly, are acoustically marked in the signal. Among these physical cues are the aspiration of word-initial stops in English (Lehiste, 1960), the lengthening of onset syllables (Gow & Gordon, 1995), as well as in some contexts the lengthening of final syllables (Beckman & Edwards, 1990). Quené (1992), for example, could demonstrate that Dutch listeners in a forced choice situation make use of durational information to word boundaries when they have to settle on one parse of an ambiguous two word utterance such as [dipɪn] either meaning *die pin* (that pin) or *diep in* (deep in). But even if some word boundaries are clearly marked in the acoustic signal this is by far not the standard situation. Most word boundaries are blurred by coarticulation with preceding and following elements. How then can the apparent ease and the effectiveness of the segmentation performance of listeners be explained?

There is now wide agreement about one central principle in segmentation: the competition process. This principle is at the core of the most influential models of speech recognition. Competition between lexical candidates occurs following the parallel activation of various entries in the lexicon that are compatible with the input. There is much evidence that lexical candidates once activated do actually compete for recognition. One of the most compelling sources of empirical support for parallel activation was supplied by a cross-modal priming study run by Zwitserlood (1989). The priming task takes advantage of the observation that structurally related words can influence one another in terms of activation. Encountering one word can either boost or inhibit the activation of other related entries (see for detailed information on priming Zwitserlood, 1997; Tabossi, 1997; Drews, 1997). A common finding in the priming literature is that lexical decision on a visually or auditorily presented word is faster to a stimulus that was preceded by a semantically related word than to a stimulus preceded by an unrelated word. The resulting RT difference is interpreted as reflecting the spreading activation that occurs between semantically related words as compared to unrelated ones. The spreading activation emerging from a certain word increases the activation of a number of related words and by that primes their recognition, which results in faster RTs to those words. Zwitserlood demonstrated that the presentation of a fragment of a word could lead to priming effects. Presenting Dutch listeners auditorily with an ambiguous string like [kapɪt] speeded their reactions to subsequently presented words like *ship* or *money*. Up to the [t] two different entries - namely [kapɪtɛin] (captain) and [kapɪta:l] (capital) - are compatible with [kapɪt]; both meanings are activated by that string and can thus both prime their semantic relatives: ship or money respectively. Although this finding convincingly demonstrated that multiple candidates can be activated in parallel it does not necessarily mean that there is actually competition between these candidate words.

Direct evidence for competition comes from various studies employing different tasks. For example, McQueen, Norris, and Cutler (1994) showed that English listeners found it much more difficult to detect words embedded in the onsets of longer words. When asked to spot words in nonsense strings, listeners were much slower to detect a target word when it was embedded in a nonsense string that itself was the onset of a longer word. Thus, spotting *mess* in [dəmɛs] - the beginning of *domestic* - was slower than when *mess* was embedded in [nəmɛs], which is not the onset of another word. The authors explained this outcome with the principle of competition: in the case of [dəmɛs] not only the target word *mess* but also the longer word *domestic* is initially activated and only at a later point in time can the recognition process favour one of the two competitors. Because [nəmɛs] is not the beginning of an existing English word the activation of *mess* is not hindered by the activation of other possible candidates, leading to shorter RTs.

Another study by the same authors (Norris, McQueen, & Cutler, 1995) supplied not only further evidence for the idea of competition between lexical elements but also provided a direct demonstration of how competition can assist the segmentation process. Participants in this study were again asked to perform a word-spotting task. One subset of items was designed such that embedded target words with the structure C(C)VCC (e.g., *stamp* or *mint*) were followed by two different kinds of nonsense syllables. One subset of items was combined with strong CVC syllables that had as initial phoneme the last one of the embedded target word (e.g., *taup* in *mintaup* with the embedded word *mint*) and that matched only **few** existing words (the few competitor condition). The other subset of items was also combined with strong CVC syllables to form nonwords (e.g., *pidge* in *stampidge* with the embedded word *stamp*) but these nonsense syllables were chosen to match **many** existing words (many competitors condition). The logic was that the spotting of target words in both conditions should be relatively hard because the second strong syllables would bias the parser to assume a word boundary *before* the onsets of the second syllables (i.e., before the last phoneme of the embedded word; this is predicted by the Metrical Segmentation Strategy as formulated by Cutler and Norris (1988) which will be discussed later in more detail). This would impair the correct segmentation and thus make the spotting of the intended words hard relative to word-spotting in strong-weak sequences (e.g., [mɪntəp] and [stæmpədʒ]). Because the parser is less likely to assume a syllable boundary before a weak syllable (see Cutler & Norris, 1988) the spotting of the words *mint* and *stamp* respectively is relatively easy as compared to the strong-strong situation. With the few-many-competitors manipulation Norris et al. (1995) wanted to demonstrate that the task would get even harder if the second strong syllable (like *pidge* in *stampidge*) activated many lexical entries, since this in turn would increase the bias to assume a word boundary before the last phoneme of the embedded target word. Norris et al.’s results were in the predicted direction: the spotting of words in the many-competitor condition (i.e., *stampidge* and *stampedge*) was on average 89 ms slower and 9% less accurate in the strong-strong contexts than in the strong-weak contexts. In the few-competitor condition (i.e., *mintaup* and *mintep*) these effects were attenuated (40 ms RT-difference and 2% error difference).

Very similar evidence about the influence of second syllable competitors on segmentation was provided by Vroomen and Gelder (1995) in Dutch. They used a cross-modal identity priming task where subjects heard words embedded in different contexts just before they had to perform a lexical decision task on the embedded words visually presented on a computer screen. Lexical decision latencies to *melk* (milk), for example, were faster when participants had just heard [mɛlkøm] with [køm] activating few competitors than when they had heard [mɛlka:m] with [ka:m] activating many competitors. Note that in both conditions the first phonemes of the second (nonsense) syllables are the last phonemes of the target word (i.e., [k] in [køm] or [ka:m] is the last phoneme of the target word *melk*). Thus, the more lexical candidates there are activated by the second syllable the harder it gets for the parser not to assume a word boundary before the onset of that syllable. Fewer competitors make it easier to identify the first phoneme of the second syllable as being the last phoneme of the target word.

Also in line with these studies is the research done by Luce et al. (1990) on neighborhood density and its influence on word recognition. For an auditory lexical decision experiment on nonwords, Luce et al. (1990) constructed nonwords with either many or few phonetic neighbors. In addition to that, the factor neighborhood frequency was manipulated resulting in four different groups of nonwords (high neighborhood density + high neighborhood frequency; high neighborhood density + low neighborhood frequency; low neighborhood density + high neighborhood frequency; low neighborhood density + low neighborhood frequency). The logic of this experiment was that the correct rejection of nonwords should vary as a function of the number of competitors and of the high or low frequency of those competitors. Thus, the more words (in a high density neighborhood) are activated, the harder it should be to reject the stimulus as being a nonword. The same holds for frequency: the higher the frequency of the neighbors the harder the task to reject the nonwords. Both factors showed significant effects on lexical decision. Participants' rejection times were significantly slower when the nonwords had many competitors as compared to rejection times for nonwords with few phonetic neighbors. Similarly, rejection times for nonwords with high frequency neighbors were slower than rejection times to nonwords with low frequency neighbors. These results support the idea that multiple activated candidate words compete during spoken word recognition (see also Luce & Large, 2001; Vitevitch & Luce, 1998, 1999).

But even if competition is empirically well supported, it still remains to be clarified how competition of lexical candidates can actually help to solve the segmentation problem in the processing of real speech, that is, outside an experimental setup like the ones described above. Assume an input that activates many possible candidates that at some point in time are compatible with that input: *unbelievable overrepresentation*. Words like *believe, believable, leave, below, over, representation* - among others - will be activated. But the words *unbelievable* and *overrepresentation* will finally win the competition process because (a) they do not inhibit each other and (b) they will provide the best and most complete parse of the input. Other words, however, will be inhibited because they cannot join forces to win over other competing candidates. Thus, a parse settling for example on the candidates *?? believe a below ?? representation* will leave the sequences [ʌn] and [ver] unaccounted for.

Even if the boundary between the intended words was not marked by clear acoustic cues the correct position of this boundary can nevertheless be correctly detected because the competition process settles on the most appropriate parse. This means that it finds the set of words that spans the whole input and not a set that spans only part of it.

Even if competition is strongly supported by theoretical considerations and empirical data, however, it is not the only device by which segmentation can be successfully achieved. There is by now a considerable body of evidence that listeners can make use of the metrical structure of their language to find boundaries between words that are not acoustically marked. Using the syllable monitoring task, Mehler, Dommergues, Frauenfelder, and Seguí (1981) and Seguí, Frauenfelder, and Mehler (1981), for example, showed that French listeners make use of the syllable as a segmentation unit. Thus, if listeners have to detect a target string such as *ba* or *bal* within a longer string like *balance* (balance; open initial syllable) or *balcon* (balcony; closed initial syllable) they need less time to do so when the string matches the syllable structure of the word than when it mismatches that structure. Thus, detecting *ba* in *balance* was faster than detecting *bal* in *balance*, and detecting *bal* in *balcon* was faster than detecting *ba* in *balcon*. Based on results from studies in other languages (see below) it has been assumed that these effects from French were not simply due to a *syllabic* segmentation strategy, but rather to a language-specific instantiation of a universal segmentation principle based on the *rhythmic* properties of each given language. This principle - the so-called Metrical Segmentation Strategy (MSS; Cutler & Norris, 1988) - proposes that listeners segment the speech according to those units that determine the rhythmic structure of their language. Because languages differ according to their metrical structure, this universal strategy should vary across languages according to the rhythmic properties of those languages.

This prediction has been confirmed by many studies. Results from Spanish and Catalan (Bradley, Sánchez-Casas, & García-Albea, 1993; Sebastian-Gallés, Dupoux, Seguí, & Mehler, 1992) - both syllable-timed languages - showed that listeners, as in the French study, used the syllable to segment the speech input. In a comparative study of French and English, Cutler, Mehler, Norris, and Seguí (1986) showed however that English listeners could not rely on the syllable as a segmentation unit. The English language has a much less regular syllable structure than French. Many intervocalic consonants in English are ambisyllabic which means that they structurally belong as much to first as to second syllables (e.g., the [l] in *balance*). As a consequence of this ambisyllabicity the authors argued that employing a syllable-based segmentation strategy in English would not only be ineffective but might even hinder the segmentation process. In order to test this hypothesis, they conducted the syllable monitoring task that Mehler et al. (1981) had used in the original French study. English listeners were asked to monitor auditorily presented words for either open syllables like *ba* or closed syllables like *bal* in two different word environments. Target-bearing word pairs were selected to share the same initial phonemes (CVC) - for example *balance-balcony*. These words differed according to the status of the third phoneme, which was either ambisyllabic, as in *balance*, or belonged clearly to the first syllable, as in *balcony*. If English listeners ignored ambisyllabicity and used a syllable-based segmentation strategy, the same cross-over effect as described by Mehler et al. (1981) for French should have been observed. This would mean that English listeners treated the ambisyllabic third phoneme (e.g., [l] in *balance*) as belonging to the second syllable. If so, finding *ba* in *balance* should have been easier than finding *bal* in *balance*. The opposite pattern should have been observed for *balcony*. Yet another possibility was that a syllable-based strategy would be found for easily syllabified words (i.e., *balcony*) but not for the other group of words (e.g., *balance*). Or the detection of CVC targets (i.e., *bal*) might overall be faster than the detection of CV targets (i.e., *ba*) because the former is present in both words in terms of syllable structure even if the [l] in *balance* also belongs to the second syllable.

None of the results that would have supported a syllabification strategy for English was found in this study. Instead, words with ambisyllabic consonants elicited faster RTs as compared to words with a clear syllable structure, irrespective of the target that had to be detected. Furthermore, English listeners were also tested on the experimental materials that had already been used in the French study by Mehler et al. (1981). As in the previously described experiment, there was no sign in the results of a segmentation strategy for English listeners based on syllabification. But when French listeners were asked to detect target strings in English materials they showed a clear syllabification strategy, just as they had done in the Mehler et al. (1981) study. Cutler et al. (1986) concluded that the syllable was not the appropriate unit to subserve the segmentation process for native English listeners, whether they listened to their own language or to a foreign language in which syllabification would have been appropriate.

In a later study, Cutler and Norris (1988) showed that English listeners use the stress information of their language for segmentation rather than the syllable. When English listeners had to detect monosyllabic words (e.g., *mint*) embedded in nonsense bisyllables, they were faster in doing so when the bisyllabic nonsense string consisted of one strong and one weak syllable (e.g., *mint* in [mɪntəf]) than when the string consisted of two strong syllables (e.g., *mint* in [mɪnteɪf]). The authors interpreted this outcome as a result of the specific segmentation strategy of English listeners. When the second syllable is strong, English listeners tend to assume a word boundary before that syllable (i.e., before the syllable [teɪf]), which in turn makes it harder to realize that the first phoneme of that syllable is actually the last one of a target word (i.e., *mint*; see also Cutler & Butterfield, 1992; Norris et al., 1995; Vroomen & Gelder, 1995; Vroomen, Zon, & Gelder, 1996).

Further evidence for the metrical segmentation strategy came from a study in Japanese, the ideal test case for the postulated principle. Otake, Hatano, Cutler, and Mehler (1993) predicted that the relevant unit for segmentation in Japanese should be the mora since this language has moraic rhythm. Otake et al. (1993) using the fragment detection task, showed indeed that Japanese listeners found it easiest to segment speech at mora boundaries. The target *ta* was detected equally rapidly in words like *tanishi* or *tanshi* because it corresponds to the first mora of both *ta-ni-shi* and *ta-n-shi*. The target *tan*, on the other hand, was hardly ever detected in *tanishi* where - in terms of moraic structure - it is simply not present. Taken together, these results provide strong evidence for a universal segmentation strategy that is based on rhythm.

The strategies described so far can be supplemented by yet another mechanism that provides listeners with segmentation information. McQueen (1998) showed that, in segmentation, listeners can use their knowledge about what phoneme sequences within a syllable are legal in their language. Using the word-spotting task, he demonstrated that Dutch listeners found it significantly harder to spot words like *rok* (skirt) in a bisyllabic nonsense string such as [fi.drok] than in a string like [fim.rok]. While the latter sequence clearly indicates a boundary before the word *rok* because [mr] is not a legal syllable onset (or coda) in Dutch, this is not the case in the string [fi.drok] where the word boundary is misaligned with the syllable boundary. Thus, it is easier for listeners to spot a word in a context where phonotactics force a boundary that is aligned with the word's onset phoneme.

Similarly, Finnish listeners can use information about the potential segmental content of a word based on vowel harmony to assist in segmentation. In Finnish certain vowel combinations are illegal within a word: the vowel in the first syllable of a longer word determines the quality of the following vowels, which have to belong either to the same vowel set or a neutral vowel set. As Suomi, McQueen, and Cutler (1997) showed, Finnish listeners find it much easier to detect a word like *käry* (odour) when it is embedded in a nonsense string where the preceding context syllable contains a vowel that violates the vowel harmony principle like in *pokäry* than in a string like *pökäry*. Because the vowels in *käry* mismatch the vowel information given in the first syllable of *pokäry* it is easy for Finnish listeners to assume a word boundary before the target word. The task gets harder when the vowel information of a longer string does conform to the vowel harmony principle, as in *pökäry*, where all the vowels belong to the front vowel set (see also Vroomen, Tuomainen, & Gelder, 1998).

Recently, Weber (2001) has demonstrated that phonotactics are not only used by listeners when they segment their native language but that listeners also apply the phonotactic constraints of their native language when listening to a second language. Furthermore, she showed that listeners with a certain grade of proficiency are able to acquire knowledge about phonotactic constraints in a foreign language and use this newly acquired knowledge to segment the non-native language. In her study, native German listeners, all of which were highly proficient in English, had to spot English words that were embedded in four different nonword environments. In the first condition, that was predicted to be the hardest, word boundaries were not clearly marked by either German or English phonotactics (i.e., *length* in [fuklɛŋθ] with [kl] being a legal cluster at both German and English word onsets). In a second condition, the embedded words' onsets were aligned with a syllable boundary in English but not in German (i.e., *length* in [zarʃlɛŋθ], with [ʃl] being a legal onset of German but not of English) while in a third condition German phonotactics forced a boundary before the word whereas English phonotactics did not (i.e., *length* in [joʊslɛŋθ] where [sl] is legal in English but not in German). In the fourth condition, the phonotactics of both languages imposed a boundary before the target words (i.e., *length* in [funlɛŋθ] with [nl] being an illegal syllable onset in either language). The results showed that German listeners were slowest in spotting English target words when their onsets were not clearly marked by either English or German phonotactics. When there was a forced syllable boundary before the embedded word, however, it made no difference whether this boundary was imposed by German or English phonotactics. German listeners were faster in both conditions than in the 'no boundary' condition, showing that they used both their native and their non-native knowledge to perform the task.

To summarize, various sources of information have been shown to be of use in solving the segmentation problem. A strong model of speech recognition should therefore try to incorporate all these findings. These different sources of information should join forces in the process of finding the correct word boundaries. The most effective model would be able to unify the effects of different cues in one single segmentation mechanism. The Possible Word Constraint (PWC), as proposed by Norris et al. (1997), provides one such unified account. According to McQueen (1998) and Norris et al. (1997), the core mechanism of segmentation is the competition process which allows the system to settle on an optimal parse even if there are no cues as to where word boundaries are. But if there are such cues, the PWC will use this information to bias and thus assist the competition process.

The Possible Word Constraint

The motivation for the formulation of the Possible Word Constraint goes back to the following line of argumentation. Norris et al. (1997) stated that the segmentation strategies found up to that time could still not explain satisfactorily how listeners can segment real speech input with such efficiency. They argued that all the strategies reviewed above had to meet two fundamental criteria in order to work: (1) the spoken input needs to be clear and unambiguous and (2) the intended word(s) have to be in the listener's vocabulary. But both criteria are very often not fulfilled in speech that is encountered in everyday life. In many situations the speech input is distorted by various factors (e.g., environmental noise) or some word might be unknown to the listener. People's performance nevertheless remains excellent in spite of these drawbacks. How, so the authors asked, can a listener easily interpret the utterance ”met a *fourf* time”even if one of the words (*fourf*) is very unlikely to be in his or her lexicon? There needs to be a mechanism that guarantees that **all** of the input is accounted for in the parse even if some words are unknown. Assume what would happen if the parser ignored the phoneme [f] because *fourf* is not listed in the lexicon: another parse would be highly activated, namely *metaphor time*. Because the embedding of words within other words or word sequences (like *metaphor* in *met a fourf*) is extremely widespread in the English vocabulary, the strategy of ignoring single consonants would lead to a severe problem of frequent misunderstandings. Every time listeners are confronted with sloppy speech there would be the danger of false and incomplete parses. Clearly this is not what we experience in normal speech processing.

Norris et al. therefore propose that human listeners can use their knowledge about the minimal size of words of their language in the segmentation of incoming speech input. This simply means that the competition process parses the speech such that no impossible words are left over. For all European languages it holds that possible words have to contain at least one vowel. As soon as it is realized that the activation of a particular lexical candidate would result in a segmentation leaving over a single consonant, the activation of this candidate is immediately reduced. Thus, the word *metaphor* will only be activated very briefly when hearing the string *met a fourf* because the selection of that word would leave the sound [f] unaccounted for.

Norris et al. (1997) tested this hypothesis experimentally. They predicted that in a word-spotting task the detection of *apple* in a nonsense string like [fæpəl] would be much harder than in a string like [vʌfæpəl], because the single consonant [f] is not and could never be an English word, whereas [vʌf] could be a word although it happens not to be. This should also hold for words in following contexts like *sea* in [siʃ] and [siʃʌb] respectively. Norris et al. (1997) reported that RTs were significantly longer when listeners spotted words in impossible contexts (like the single consonant [f] in [fæpəl]) than when listeners spotted words in possible contexts (e.g., *apple* in [vʌfæpəl]). Furthermore, error rates were higher when listeners had to spot words in a single consonant context ([fæpəl]) as compared to the *syllabic* context ([vʌfæpəl]). The authors therefore concluded that the PWC restricts the activation and competition process to disfavour candidate words not aligned with possible word boundaries.

The great advantage of a principle like the Possible Word Constraint lies in the integration of different cues the constraint can make use of to locate word boundaries. It does not depend only on phonotactic or only on rhythmic information, but can use both of these sources of information when they are available as well as other cues to the location of likely word boundaries. If not cued by information in the signal, word boundaries nevertheless will be correctly found because the competition process together with the knowledge of the necessary size of words provides a mechanism to locate such boundaries. By implementing the PWC in the Shortlist model, Norris et al. (1997) demonstrated that the segmentation process was indeed more effective when the competition process was biased by the PWC than when it was not.

Languages, however, not only differ in their metrical structure - as already outlined above - but also differ in terms of what constitutes a possible word in a given language. A legal word in French, for example, can consist of an open syllable with a lax vowel like *thé* (tea); this is not the case in English. Here a legal word with a lax vowel has to be a closed syllable (like *book*); open syllables have to have a tense vowel (like *sea*). An open syllable like [bʊ] with the short vowel from *book* is not legal in English. In yet other languages like the Bantu language Sesotho, however, a content word has to consist of at least two syllables.

The question thus arises whether the PWC is sensitive to this language-specific variation. Norris, McQueen, Cutler, Butterfield, and Kearns (2001) investigated exactly that question in English. They predicted that if the PWC was language-specific, English listeners should find it just as hard to spot *canal* in either [skənæl] or [zɛkənæl] because both residues [s] and [zɛ] are not legal words of English. Note that open syllables with a lax vowel like [zɛ] are not permissible in English. Spotting canal in [ziːkənæl] (with [ziː] being a legal English syllable), on the other hand, should be easier than in either of the other two contexts. If, however, the residue only needs to be a possible word of any language - implying the universality of the constraint - spotting *canal* in [zɛkənæl] should be as easy as in [ziːkənæl]. Although [zɛ] is not a possible English word it could well be a word in French, for example. Thus, if the PWC is satisfied by any word of any language, the activation of *canal* in [zɛkənæl] should not be reduced. Its detection should therefore be easier in [zɛkənæl] than in [skənæl]. The results showed that English listeners' performance was as good if the residue was an illegal syllable of their language as when the residue was a legal syllable. Thus, word spotting latencies in both conditions (i.e., [zɛkənæl] and [ziːkənæl]) were faster than in the *consonantal* condition (i.e., [skənæl]). This finding was further supported by a second experiment showing that the spotting of a word in a CəC context (i.e., *sea* in [siʃəb]) was easier than the spotting of a word in *consonantal* context (i.e., [siʃ]) even though in English only function words can consist of a single weak syllable.

Closely related to that study was an experiment conducted in Dutch by McQueen and Cutler (1998). Three conditions similar to those in the English experiment were tested. But in addition to that a fourth condition was introduced in order to check whether the PWC works in graded fashion, that is, whether some possible words are better than others in terms of segmentation. Thus, four different nonword contexts were contrasted: words were either embedded in strong (CV; e.g., *lepel* 'spoon' in [kʏleːpəl]), weak (Cə; e.g., *lepel* in [səleːpəl]) or bisyllabic (CVCə; e.g., *wonen* 'to live' in [dʏkəwoːnən]) possible-word contexts, or in an impossible word context (i.e., *wonen* in [dwoːnən]). In line with the PWC, the results showed that listeners were reliably faster and more accurate in spotting words that were embedded in possible-word contexts than when they had to spot words in impossible word (i.e., single consonant) contexts. Furthermore, it was demonstrated that all possible word contexts were equally acceptable as segmentation cues for the PWC. The RT difference that the authors found between the CV and the Cə contexts (with the latter being faster than the former) could be attributed to acoustic differences between the embedded target words, as was established in a later control lexical decision experiment.

Together with the results from the English experiments described earlier, these findings convincingly demonstrate that the PWC is not sensitive to language specific information about the minimal size of words but instead that the PWC is satisfied by any syllable. Recent results from word-spotting studies in two non-European languages further supported this universality claim. McQueen, Otake, and Cutler (2001) report that for Japanese listeners, segmentation performance was as good when the residue of a given sequence consisted of a single vowel as when it consisted of a single moraic nasal consonant (both possible morae -

the relevant segmentation unit - of Japanese). Thus, Japanese listeners detected *saru* ’monkey’, for example, as fast in *sarua* (vowel context) as in *saruN* (moraic nasal context) because in both cases the word is aligned with a metrical boundary. If, however, the residue only consisted of a single nonmoraic consonant (i.e., *saru* in *sarup*) listeners found it much harder or - in some circumstances - impossible to detect an embedded word. Thus, the authors concluded, these results not only confirm earlier findings on the moraic segmentation strategy used by Japanese listeners, they also shed more light on the generality of the PWC.

The parser appears to process the incoming speech on the basis of the PWC - regardless of what language is spoken - such that no chunks only consisting of consonants are left over. Whether a single vowel is a ”possible word”in a given language or not is irrelevant for the working of the PWC. As long as each chunk of speech in a given parse of the input contains at least a vowel, the activation of candidate words in the parse will not be penalized. As soon as the parser encounters a single consonant or a consonant cluster as the residue in a given segmentation, however, the activation of a candidate word will be penalized. Recent results from a word-spotting study in Sesotho, a South African language in which a content word needs to be minimally bisyllabic, are in line with this argument. As Cutler, Demuth, and McQueen (submitted) report, Sesotho listeners find it as easy to spot words in monosyllabic (CV) as in bisyllabic (CVCV) contexts although the former cannot constitute a content word in that language. This again suggests that the PWC can be satisfied by a syllable that consists of at least one vowel. Even though a CV syllable is not a possible word in Sesotho, it nevertheless passed the PWC.

To summarize, the PWC appears to be a mechanism that incorporates in a universal fashion various segmentation strategies. If available, acoustic, metrical or phonotactic information will signal likely word boundaries and these boundaries will be used to bias the activation of those candidate words that are in the competition process. These acoustic, phonotactic, and metrical cues are subject to language-specific phonological restrictions. In addition to these cues, the PWC provides a further source of segmentation information that influences the competition process. This information appears to be universal. It signals to the parser that a word is a plausible candidate only if the stretch of speech between the edge of that word and a likely boundary contains at least a vowel. Whether or not a single vowel is sufficient to form a word in a given language is not relevant for the operation of the PWC. Results so far suggest that the PWC is a purely phonological mechanism that is not sensitive to higher-order knowledge. It seems to be clear that the parser locates word boundaries to avoid stretches of speech that only consist of consonants.

The case of single consonant morphemes

In some European languages, like German, Dutch, or English, one single consonant cannot constitute a word but nevertheless can be a meaningful unit in morphological terms (e.g., 3^{rd} person singular *-t* in a Dutch inflected verb like *loop+t*, 'he/she walks'). What implications does the PWC have for the processing of these inflected words? As was already mentioned in Chapter 1, there are two extreme positions as to how morphologically complex words are represented at the lexical level: there is the full-listing account which states that each complex word has its own full-form representation and there is the decompositional account which assumes separate representations for each morpheme.

If inflected complex words are in fact decomposed during segmentation, then form representations for affixes are needed for the mapping process. A form like *loop+t* (he/she walks) would thus be decomposed into the verbal stem *loop*-$_V$ and the 3^{rd} person singular suffix *-t*, both of which would need to have separate access representations. If, however, there is no decomposition of inflected words, that is, if decomposition does not take place during segmentation, then *loop*-$_V$ and *-t* would not need to have separate access representations.

The consequences for the PWC are very clear cut. If there is indeed decomposition of complex forms prior to lexical access, then the PWC has to be sensitive to morphological information in order to guarantee a smooth segmentation process. Otherwise many inflected forms would hinder the segmentation process in the following way. If the parser came across the decomposed parts of a word like *loop+t* and if the PWC were not sensitive to morphological information, the activation of the correct stem *loop* would be penalized because a single consonant like *-t* would fail the PWC. This in turn would lead to a reduction of the activation of the correct word *loopt*. Thus, decomposition requires the PWC to be morphologically sensitive. This line of reasoning also directly implies that if the PWC is *not* sensitive to morphological information, there need to be full form access representations of inflected words so that the correct inflected forms can enter the competition process for recognition.

Experiment 2.1A: Word-Spotting

The first experiment is dedicated to the investigation of how inflectional information is mapped onto meaning during spoken language comprehension. As already discussed in Chapter 1, the transition of acoustic information to meaning representations is supposed to be mediated by *access* representations that are modality dependent. The specific question of this study is therefore: are inflectional morphemes already realized as such during the segmentation of spoken language, that is, are there individual access representations for inflectional morphemes? Will listeners still employ the PWC when confronted (experimentally) with illegal combinations of morphemes like for example in [dø:rt] (i.e., **deur*$+$*t* ’door$+$t’) where a verbal inflectional marker is incorrectly joined to a noun? Would these morphemes during the segmentation process be treated like ”impossible words”and therefore fail the PWC, or, given the immense syntactic power and the high frequency of inflectional morphemes, would they have the same status for the PWC as ”real”words and therefore pass the PWC?

The first experiment examined Dutch listeners’ ability to spot monosyllabic nouns, like *deur* ’door’, in a nonsense string like [dø:rt] (*morphological* context; *-t* is the 3^{rd} person singular marker for verbs in Dutch) as compared to a *consonantal* context [dø:rp] (where *p* is not a morpheme of Dutch) and a *syllabic* context [dø:rtɑχ]. If affixes fail the PWC (i.e., are not perceived as any different from other *meaningless* consonants), then the detection of *deur* in [dø:rt] should be just as hard as the detection of *deur* in [dø:rp]. This would imply that the PWC is a purely phonological mechanism, that is, one not sensitive to morphological information. Furthermore, it would suggest that complex inflectional forms like *loop+t* are not decomposed prior to lexical access, and thus that the complex form would actually have its own access representation. The inflectional suffix *-t* would thus not cause the penalization of the complex form *loopt*, but would nevertheless cause, by triggering the PWC, a suppression of the form *loop* as a competition candidate. If, however, inflectional affixes pass the PWC, then the detection of a word embedded in a *morphological* context (i.e., *deur* in [dø:rt]) would be easier and therefore faster than the detection of a word embedded in a *consonantal* context (i.e., *deur* in [dø:rp]). Following the former line of argumentation, this would show that the PWC is influenced by morphological information. This operation of the PWC would be consistent with decomposition prior to lexical access, implying that the activation of the stem *loop-*$_V$ would not be penalized by the inflectional suffix *-t* because the parser would be able to detect a ’word boundary’ between the stem and the suffix. This latter result would thus suggest that listeners use their knowledge about inflectional morphology while segmenting acoustic speech input.

Method

Materials. The stimuli of the first experiment were based on a set of 30 Dutch words all of which were monosyllabic. All words were Dutch nouns with only two exceptions: *nul* ’zero’ and *vier* ’four’. Half of these words were embedded in nonsense strings by the addition of *-t* (marker for 3^{rd} person singular in Dutch verbs; e.g., *deur* ’door’ in *deurt*) while the other half were combined with *-s* (one of the two plural affixes for Dutch nouns; e.g., *duim* ’thumb’ in *duims*; henceforth these will be referred to as *morphological* contexts). Note that all the resulting items were nonwords in Dutch: the nouns chosen for this experiment do not take verbal affixes, and all the nouns chosen take *-en* as the plural affix, **duims* for example is thus the wrong plural form of *duim*. It is clear that the combination of nouns with the two different morphemes leads to very different linguistic ’violations’. The combination of a noun with the inflectional marker *-t* is a legal operation in Dutch which occurs quite frequently (e.g., *vis*$_N$ ’fish’ can become *vist*$_V$ ’she/he fishes’). But all nouns in the present study are usually not used as verbs in the Dutch language. The combination of a noun with the plural marker *-s*, on the other hand, results in a violation of an already existing correct plural form (*duimen* for example being the correct plural of *duim*). Thus, it is possible that the results for the two different morphemes would be different. The same words that were used in the *morphological* contexts were also embedded in nonsense strings by the addition of the consonants *p, k,* and *f*. These are not morphemes of Dutch (e.g., *deurp; duimf*; these will be referred to as *consonantal* contexts). In a third condition, the *possible-word* condition, all targets were embedded in bisyllabic nonsense strings with the added syllable always being a strong syllable. All second syllables began with one of the five consonants used in the *morphological* and *consonantal* contexts (i.e., [t,s,p,k,fVC]; e.g., *deurtach*). Both target words and following consonants were controlled for phonotactic features in order to avoid clusters that were illegal in Dutch. Otherwise subjects might have been able to use phonotactic information to solve the experimental task (see McQueen, 1998). For example, if listeners had to spot *ding* ’thing’ in a nonsense string like [dɪŋp] that ends in an illegal consonant cluster (i.e., *[ŋp]), they would be able to use the phonotactic information for the word-spotting task as well as the PWC.

Those items that were combined with *-t* in the *morphological* context were also presented in combination with either *p* or *k*, whereas items combined with *-s* in the *morphological* condition were presented with an *f* in the *consonantal* context. The phonemes in the *consonantal* condition were chosen because they are phonologically as similar to the suffixes *-t* and *-s* as possible ([t, k, p] are voiceless plosives that differ only in place of articulation, and [s, f] are both voiceless fricatives that again only differ in place of articulation). It was not possible to have only *k* or only *p* in the *consonantal* condition because some of the embedded nouns could either be combined with *k* but not with *p* (e.g., *wang* ’cheek’ in [wɑŋk] vs. *[wɑŋp]) or the other way round (e.g., *rum* ’rum’ in [rʏmp] vs. *[rʏmk]). Therefore two different control consonants had to be chosen. In order to balance the number of mono- and bisyllabic nonsense strings containing embedded words, 10 extra target-bearing items were added. These were constructed in the same way as the other bisyllabic items, the only difference being that they contained words that were not experimental targets. All nouns consisted of a [CVC] cluster with three exceptions that were [CCVC] clusters and one that was a [CVCC] cluster. The nonsense strings did not contain any other words than the target items. See Appendix A for a full list of items.

In addition to the 100 target-bearing items (including the 10 extra items) there were 80 filler items without any embedded words. These were constructed along similar lines to the experimental items, and consisted of 40 monosyllabic and 40 bisyllabic strings. 30 of the monosyllabic filler items contained the same proportion of ”affixes”and meaningless consonants as the experimental items, while 10 ended with other consonants. The bisyllabic fillers had the same structure as the experimental bisyllabic items. The materials were divided into three subsets such that each list contained 10 items in each condition and such that all target words appeared in all subsets in the same positions. The only difference between these subsets was in terms of which context a given target was presented in. Two differently randomized versions of each of these lists were constructed. Thus, within each subset the two lists contained the same experimental target-bearing strings (10 in each condition) but the order of presentation differed between the lists. This method produced a total of six different item lists. Subjects were presented with only one of these lists so that they heard each target only once.

Recording. Materials were read by a female native speaker of Dutch onto Digital Audio Tape (DAT) in a sound attenuated booth. After stimuli were redigitized onto a computer, the materials were measured and spliced into individual speech files, using the Xwaves/ESPS speech editor. For use in the experiment the separate speech files were then transferred to a PC.

Subjects. Thirty-six Dutch native speakers participated in the experiment, 30 of which were female and 6 were male students of Nijmegen University. None of them had any known hearing disorder. They were paid for their participation.

Procedure. The presentation of stimuli as well as the recording of the manual responses and the RTs were performed by the NESU software that was developed at the Max-Planck-Institute for Psycholinguistics. For later inspection of the subjects’ responses, all vocal reactions were recorded onto Digital Audio Tape. Stimuli were presented over Sennheiser headphones and each participant was

Table 2.1:
Experiment 2.1A: Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions.

	context		
	morphological	consonantal	syllabic
RT (ms)	885	932	853
Error rate (%)	19	21	10
Examples	deurt/duims	deurp/duimf	deurtach/duimfoel

tested in an individual sound-attenuated booth. Before the experiment, subjects were given written instructions explaining that they had to press a button on a box in front of them whenever they heard a nonsense string containing a real word. They were asked to speak out aloud the word they had spotted after pressing the button. There was a short break between the practice and the experimental sections. No pause was required during the experimental session, which lasted less than 15 minutes. After each stimulus there was a standard pause of 3500 ms before the next item was presented.

Results

Raw reaction times, measured from item onsets, had to be adjusted by subtracting the lengths of the embedded word in order to yield RTs from word offsets. Responses where subjects pressed a button but either failed to make a vocal response at all or spotted the wrong word were treated as errors and were thus excluded from further RT analyses. This was true for a proportion of 2.2% of all responses. One subject had to be excluded from further analyses because he failed to identify any target combined with *-t* in the *morphological* context.

The item *rum* ’rum’ was not included in any analysis because it produced an overall error rate of 47% in the present experiment. In addition to that, in a Lexical Decision control experiment (see Experiment 2.1B below) the same item was recognized less than 50% of the time. All other items in the current experiment as well as in the control experiment produced less than 45% errors and were kept in the analyses. Table 2.1 and Figure 2.1 show the mean RTs and error rates per experimental condition.

Analyses of variance (ANOVAs) were calculated for both RTs and error rates with both subjects (F_1) and items (F_2) as the repeated measures. In the RT analysis there was a main effect of context type that was only significant in the subjects’ analysis ($F_1(2,64) = 4.62$, $p < .05$; $F_2(2,54) = 2.21$, n.s.). There was

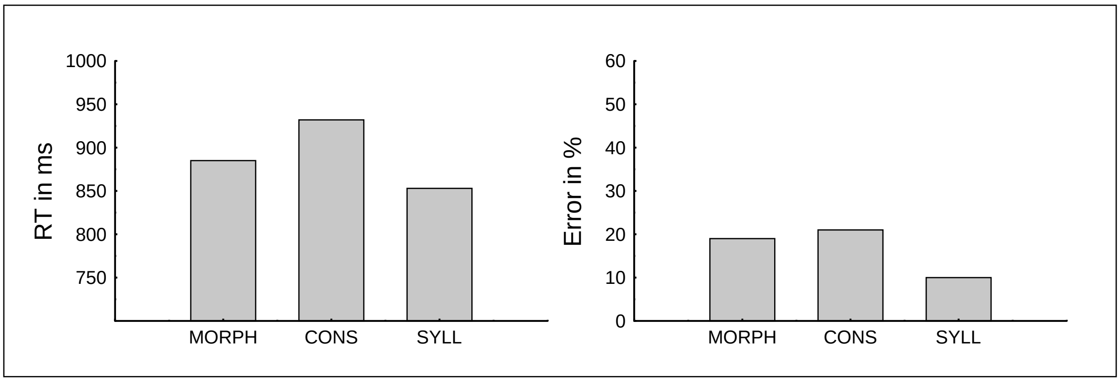


Figure 2.1: Experiment 2.1A: Mean RTs (in ms) and mean error rates (in %) per experimental condition (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

also an effect of the factor phoneme (whether the morpheme and the matched consonants were stops or fricatives) that was again only reliable in the subjects' analysis (F_1 (1,32) = 4.72, $p < .05$; F_2(1,27) = 2.39, n.s.). The interaction of these two factors was not significant. In the error analysis there was a main effect of context type (F_1(2,64) = 7.98, $p < .01$; F_2(2,54) = 4.38, $p < .05$) as well as of the factor phoneme (F_1(1,32) = 28.27, $p < .000$; F_2(1,27) = 7.43, $p < .01$) while the interaction of these factors was again not significant.

Responses to targets in the *consonantal* contexts were 79 ms slower than responses to targets embedded in *syllabic* contexts. As an overall t-test showed, this difference was significant by subjects and by items (t_1(34) = 3.01, $p < .01$; t_2(28) = 2.36, $p < .05$). Although RTs to targets in the *morphological* condition were on average 47 ms faster than in the *consonantal* and 32 ms slower than in the *syllabic* condition these differences were not significant. A somewhat different pattern was found for the error analysis. Listeners made significantly more errors when they spotted words in *consonantal* than when they spotted words in *syllabic* contexts (t_1(34) = 3.58, $p < .01$; t_2(28) = 3.06, $p < .01$). In contrast to the RT data, listeners' performance in the *morphological* condition was now significantly different from the *syllabic* condition (t_1(34) = 3.06, $p < .01$; t_2(28) = 2.11, $p < .05$) but not from the *consonantal* condition. RTs and error rates were then analysed separately for the different sets of phonemes (henceforth referred to as *stop* vs. *fricative* items). The respective RTs and error rates are shown in Table 2.2 and are illustrated in Figure 2.2. ANOVAs and t-tests were then performed on both RTs and errors for the stop and the fricative items separately.

Table 2.2:
Experiment 2.1A: Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions listed separately for each phoneme type.

stop item set			
		context	
	morphological	consonantal	syllabic
RT (ms)	911	1010	853
Error rate (%)	26	29	13
Example	deurt	deurp	deurtach
fricative item set			
		context	
	morphological	consonantal	syllabic
RT (ms)	865	873	854
Error rate (%)	13	14	7
Example	duims	duimf	duimfoel

Stop items

There was a marginal effect of Context Type for both RTs and error rates (RTs: $F_1(2,64) = 4.99$, $p < .01$; $F_2(2,26) = 2.45$, $p < .1$; error rates: $F_1(2,64) = 8.61$, $p < .000$; $F_2(2,26) = 2.78$, $p < .1$). T-tests showed that items embedded in *syllabic* contexts were spotted significantly faster (157 ms) than the same items embedded in *consonantal* contexts ($t_1(34) = 3.22$, $p < .01$; $t_2(13) = 3.04$, $p < .01$). The *morphological* condition again failed to show any significant effects as compared to the other two conditions although it was 99 ms faster than the *consonantal* condition and 58 ms slower than the *syllabic* condition. The pattern for the error rates looked very similar to the overall error pattern. The *consonantal* condition showed significantly higher error rates as compared to the *syllabic* condition ($t_1(34) = 3.82$, $p < .01$; $t_2(13) = 2.88$, $p < .01$). The same tendency was observed for the comparison of the *morphological* and *syllabic* conditions but this difference was only significant in the subjects' analysis ($t_1(34) = 3.30$, $p < .01$; $t_2(13) = 1.61$, $p < .1$). There was no difference between the *morphological* and the *consonantal* conditions.

Fricative items

The ANOVA showed no significant main effects, for either RTs or error rates. There were also no significant differences in pairwise t-tests for RTs between the three conditions. The same was true for the error analysis although there

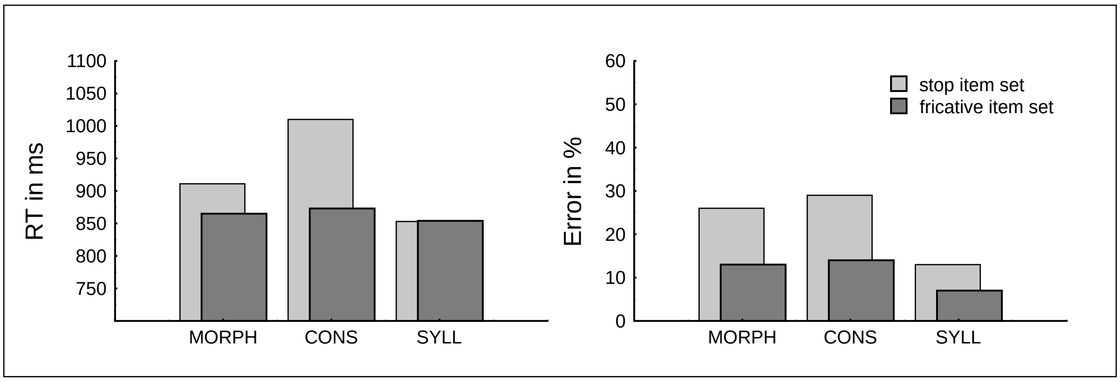


Figure 2.2: Experiment 2.1A: Mean RTs (in ms) and mean error rates (in %) per experimental condition plotted separately for each phoneme type (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

was a trend in error rates: responses in the *consonantal* condition were less accurate than those in the *syllabic* condition ($t_1(34) = 1.99$, $p < .1$; $t_2(14) = 1.41$, n.s.).

Durational Analyses

There were significant negative correlations of both RTs and errors with the lengths of the embedded words. These correlations were, however, distributed differently over the three conditions: while the RT with word length correlation was significant in the *consonantal* condition ($r(29) = -.50$, $p < .01$) the correlation of error rates with word length was significant in the *syllabic* condition ($r(29) = -.44$, $p < .05$). The mean lengths of embedded target words in the three different conditions are listed in Table 2. Generally speaking, the longer the embedded target words were, the shorter the RTs and the lower the error rates. Therefore, Analyses of Covariance (ANCOVAs) were performed across items, with the factor word length as a covariate for both RTs and error rates. These were calculated for stop items and fricative items separately because the number of cases differed between the two item sets. Neither of the ANCOVAs revealed any significant results. The marginal context effect that was observed for the error rates in the stop item set thus vanished when word duration was a covariate.

Because the lengths of the three different following contexts varied between the conditions (single consonants are shorter than syllables) correlations of RTs and error rates with the factor context-length were performed individually for each condition. None of these correlational analyses revealed a significant effect. RTs and error rates were thus not determined by the lengths of the following contexts.

Table 2.3:
Experiment 2.1A: Mean lengths of embedded target words (in ms) and mean context lengths (in ms) in each of the three conditions.

	context		
	morphological	consonantal	syllabic
Mean length (ms)	340	335	343
Mean context length (ms)	292	250	585
Examples	deurt/duims	deurp/duimf	deurtach/duimfoel

Discussion

As was predicted, these results replicate former findings from word-spotting experiments (e.g., Norris et al., 1997; McQueen, 1998). The spotting of words with possible (i.e., *syllabic*) contexts was faster and more accurate than the spotting of words with impossible (i.e., *consonantal*) contexts (although the RT effect was only reliable by subjects, the accuracy effect was fully reliable). Although the results from the error-analyses suggest that the *morphological* case behaves more like the *consonantal* condition, the pattern of RT data on the contrary suggests a somewhat intermediate status. This observation, however, is not reliably supported by the data because the *morphological* condition is not significantly different from either of the other two conditions. It is thus hard to tell whether morphemes were actually treated more like single consonants or more like syllables (i.e., like potential words) during spoken word recognition.

It was mentioned before that the morphemes *-t* and *-s* could lead to different results in the current experiment because of their different morphological status. Although the context effect did not interact with item set, a closer inspection of the data suggested that the pattern of results was indeed different for the two item sets. Items combined with *-t* show a much clearer pattern as compared to the items combined with *-s*. For the *stop* items, subjects were significantly faster to spot words embedded in *syllabic* contexts than they were to spot words in *consonantal* contexts. This is further supported by the pattern of the error rates which showed that words were spotted more accurately in *syllabic* than in *consonantal* contexts. That this result is only marginally significant can be attributed to the fact that the data set with only 14 *stop* items included in the analysis was very small and that there was high variance in the items. This marginal effect vanished, however, when word duration was added as a covariate to the analysis. This suggests that the pattern of error rates was to some extent influenced by the durations of words in at least two of the three conditions. Furthermore the analysis of the errors suggests that the morpheme *-t* is apprehended more like non-morphological consonants during auditory speech perception - the error rates for the *morphological* condition were higher than those for the *syllabic* condition but not different from the error rates in the *consonantal* condition.

Looking at the morpheme *-s* the picture becomes less clear. Although the RT pattern for those items looks like that for the *stop* items, there are no reliable differences between the three contexts. Even the well established PWC effect is absent. Thus people seem to be as fast and as accurate in spotting words in *consonantal* as in *syllabic* contexts. One reason for this outcome could be the very low sequential probabilities within the consonant clusters at the ends of the *fricative* items in the *consonantal* condition. Although all item-final clusters were legal clusters of Dutch, those in the *consonantal* condition of the *fricative* item set were very infrequent. The cluster *mf* at the end of words only occurs in loan words like *nymph* 'nymph' that by themselves are very infrequent (all frequencies are based on the CELEX computerised database of Dutch, Baayen, Piepenbrock, & Rijn, 1993; frequency of the coda cluster [mf]: 11 word types). Furthermore, the VCC phoneme sequence [irf] that occurred in three of the four items ending in [rf] is also quite infrequent since it only occurs in the past tense forms of a couple of Dutch verbs (e.g., *stierf* 'died'; frequency of [irf]: 17 word types).

The missing PWC effect in this item set could thus be due to the listeners' sensitivity to the sequential probabilities within the last consonant clusters. It has been shown by van der Lugt (1999, 2001) that listeners can use not only knowledge about the illegality of certain sequences for segmentation but that they are also sensitive to the likelihood of phoneme sequences in segmentation. That is, the more likely a sequence is to occur in a language, the more difficult it will be for the listener to assume a word boundary within that sequence. Apparently, listeners in the present study could place the word boundaries easily in the correct position because transitional probabilities indicated that the sequences of phonemes spanning the ends of the target words and the following contexts were unlikely to span word boundaries. This means that listeners in this experiment used their implicit knowledge about how likely a phoneme sequence is in their language in order to solve the task. Because the phoneme sequences used in the *fricative* item set were not very familiar to them it was fairly easy for them to assume the word boundaries at the correct positions, namely before the last phonemes in the *consonantal* condition.

Experiment 2.1B: Lexical Decision

Because all items were naturally uttered by the speaker it is possible that the results are attributable to differences in the acoustic realizations of the target words rather than to the influence of the phonemic contexts they were embedded in. It might be that target words embedded in *syllabic* contexts sounded more natural than the same words uttered in *consonantal* or *morphological* contexts. If so, the results would reflect difficulties of identification of target words due to acoustic confounds rather than a direct influence of the context itself. Therefore, as in previous studies (see McQueen, 1996), a lexical decision control experiment was conducted in order to rule out this alternative explanation.

Method

Materials. The set of stimuli was based on the same items used in the previous word-spotting experiment. Words were excised from the three different contexts, using the Xwaves/ESPS speech editor, leaving three different versions of each target word. Because there were now only monosyllabic items, the bisyllabic fillers were also all reduced to monosyllables. In order to make the fillers structurally as similar as possible to the target words, the last phonemes of the monosyllabic fillers were also cut off. All cuts were made at zero crossings based on visual and auditory inspection of the waveforms. This left a list of 100 words and 80 nonwords. The items were presented in exactly the same randomized orders as in the word-spotting experiment. Again each subject was presented with only one of the six lists so that she/he heard each target word only once.

Subjects and Procedure. Again 36 subjects participated in the experiment; 26 female and 10 male students from Nijmegen University were paid for their participation. The technical setup was identical to the one in the previous experiment. Before the experiment, subjects received written instructions explaining that they had to press a button on a box in front of them only when the item they had heard was an existing word of Dutch (a go/no-go procedure).

Results and Discussion

As in the word-spotting experiment, the raw reaction times had to be adjusted by subtracting the lengths of the words in order to yield RTs from word offsets. As mentioned earlier, one item was excluded from all of the analyses because it produced an error rate of more than 50%. Table 2.4 and Figure 2.3 show the mean RTs and error rates per condition. Table 2.5 shows the mean RTs and error

Table 2.4:
Experiment 2.1B: Mean lexical-decision latencies (RT, in ms) and Error Rates (%) in each of the three conditions.

	context		
	morphological	consonantal	syllabic
RT (ms)	457	470	475
Error rate (%)	12	12	11
Examples	deur(t)/duim(s)	deur(p)/duim(f)	deur(tach)/duim(foel)

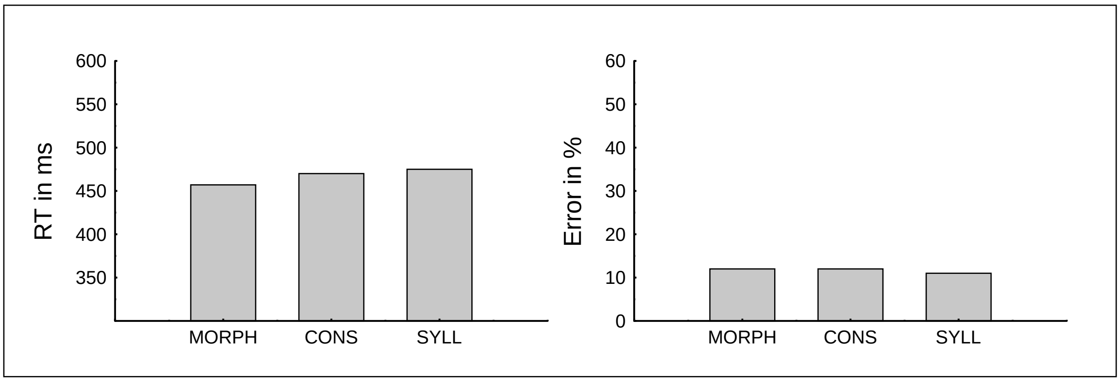


Figure 2.3: Experiment 2.1B: Mean RTs (in ms) and mean error rates (in %) per experimental condition (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

rates for each context splitting the conditions according to the set of phonemes (*stops* vs. *fricatives*) that had previously been used as contexts.

ANOVAs were performed on RTs and error rates with both subjects (F_1) and items (F_2) as the repeated measures. In both the RT and error analyses there was neither an effect of context type nor an effect of the factor phoneme. There was also no interaction of these factors, neither for RTs nor for error rates.

Although neither the factor phoneme nor the interaction of both factors were significant in any of the analyses, the data was analysed separately for each type of phoneme in order compare the data to the results of Experiment 1. Individual ANOVAs for the two item sets did not reveal any significant effects. Overall the results of the follow-up experiment convincingly suggest that the effects that were found in the word-spotting experiment are in fact attributable to the manipulation of the contexts in which target words were presented since the effects vanished after the contexts had been cut off these items. Thus, there were no differences in the realization of the embedded targets that could have been responsible for the pattern of results obtained in the word-spotting experiment.

Table 2.5:
Experiment 2.1B: Mean lexical-decision latencies (RT, in ms) and Error Rates (%) in each of the three conditions listed separately for each phoneme type.

stop item set			
		context	
	morphological	consonantal	syllabic
RT (ms)	457	462	474
Error rate (%)	11	15	13
Example	deur(t)	deur(p)	deur(tach)
fricative item set			
		context	
	morphological	consonantal	syllabic
RT (ms)	456	476	476
Error rate (%)	13	10	10
Example	duim(s)	duim(f)	duim(foel)

To summarize, the present experiments replicate former findings from word-spotting (e.g., Norris et al., 1997; McQueen, 1998) where words embedded in *syllabic* contexts elicited faster word-spotting latencies and lower error rates than words embedded in *consonantal* contexts. Thus, spotting *deur* in *deurtach* was easier than spotting *deur* in *deurp*. This difference went away in the lexical decision experiment where the embedded words were presented without the following contexts. However, the current results could not unambiguously answer the main question about whether morphemes have a special status for the PWC or not. On the one hand, there is a hint in the error data that they might not be different from 'meaningless' consonants, but the RTs, on the other hand, suggest that morphemes differ from both syllables and consonants. That the differences in RTs between the three conditions were not statistically reliable might be due to two main factors. First, the data set was quite small and thus probably not powerful enough to reveal more subtle differences. And second, there was the confound of sequential probabilities in the *fricative* item set that most likely prevented the PWC from showing its effect. It therefore seemed worthwhile to investigate the topic further while taking these confounds into account. New experiments with new materials are reported in Chapter 3.

The PWC and morphology - further investigations

Chapter 3

The question about the status of inflectional morphemes for the PWC still needs to be answered. Is the parser aware of the morphological status of inflectional morphemes during segmentation? If so, the PWC would not be a purely phonological device but would be sensitive to morphological information. This question was already addressed in Experiment 2.1A but could not be answered satisfactorily: although the error data of Experiment 2.1A seems to suggest that single-consonant morphemes are not different from morphologically meaningless consonants in terms of the segmentation principles of the PWC, the RT data indicates a more intermediate status for morphemes. Thus, while words embedded in a possible-word (i.e., *syllabic*) context were spotted reliably faster and more accurately than the same targets embedded in an impossible-word (i.e., *consonantal*) context, words in a *morphological* context elicited faster RTs as compared to the *consonantal* condition but slower RTs as compared to the *syllabic* condition. However, this intermediate status for morphemes was not supported by the statistics. Therefore a new experiment was designed in which more target words were presented in the three relevant contexts (i.e., *deur* in *deurt, deurp* and *deurtach*). Remember that the PWC predicts that a target word followed by a legal syllable (e.g., *deur* in *deurtach*) should be detectable more easily by the listener than the same target followed by a single consonant (e.g., *deur* in *deurp*). The logic is that the parser modulates the activation of words by checking the effect of a given word candidate in the context: only the activation of those words is maintained which do not leave over impossible words between word boundaries. Because single consonants are not legal words in Dutch, the spotting of *deur* in *deurp* should be slower and less accurate than the spotting of *deur* in *deurtach*. Although the syllable *tach* is meaningless in Dutch, it passes the PWC because it is a potential word. What about morphemes? If the PWC is sensitive to morphological information, the spotting of *deur* in *deurt* should be as easy as in *deurtach* since ’-t’ is an inflectional morpheme in Dutch. If, however, the PWC is not morphologically sensitive, the spotting of *deur* in *deurt* should be as difficult as in *deurp*.

In order to resolve this question, new experimental stimuli were constructed. The difference from Experiment 2.1A was that the two item sets - *stop* and *fricative* items (e.g., *deurt* and *duims* respectively) - were presented in two independent experiments. This was motivated by the results of Experiment 2.1A, which indicated a difference between these two item sets. Furthermore, the number of stimuli per item set could be extended and thus the experimental power could be increased.

Experiment 3.1

Method

Materials. No more monosyllabic words could be found that met all of the factors that had to be controlled for. Therefore, bisyllabic nouns were now included in the item lists. For phonotactic reasons, all of these new target words had the stress pattern weak-strong (e.g., *forel* ’trout’): in Dutch only full syllables can end with complex consonant clusters such as were required for the *consonantal* condition.

As already mentioned, many nouns in Dutch can be easily verbalized by adding the inflectional marker *-t* (e.g., *vis* ’fish’ in *vist* ’she/he fishes’). This severely reduced the list of candidate words that could be combined with the affix *-t* without forming a word. Thus, the number of items in the *stop* item set was necessarily lower than that in the *fricative* item set. For the *stop* item set, 21 new targets could be found that met all the constraints that had to be controlled for. From the monosyllabic item list of Experiment 2.1A, the 9 *stop* items that produced the lowest error rates were selected. This resulted in a total of 30 targets for the *stop* item set.

In the *fricative* item set, 24 new words were found which were included in the new set together with 24 monosyllabic targets from Experiment 2.1A. Because only 15 monosyllabic words had been in the earlier *fricative* item set, 9 words of the *stop* item set were now chosen to be combined with the affix *-s* in order to balance the number of mono- and bisyllabic target words.

Each of these target words was then embedded in the three types of contexts: *morphological* (e.g., *deur* in *deurt*), *consonantal* (e.g., *deur* in *deurp*) and *syllabic* (e.g., *deur* in *deurtach*). Nonwords in the *stop* item set for each of these three types of contexts were created along the same lines as in Experiment 2.1A. Nonwords in the *consonantal* condition of the *fricative* item set, however, were now constructed as follows. Those 25 nouns ending with the phoneme [l] were combined with the consonant [f], like *forel* ’trout’ in *forelf*. As reported in Chapter 2, the consonant clusters *rf* and *mf* in coda position are very infrequent in Dutch. Listeners in Experiment 2.1A seemed to be sensitive to that fact, because words of the *fricative* item set were not spotted more slowly or less accurately in the *consonantal* condition than in the *syllabic* condition. In order to avoid this phonotactic confound, the 7 words in this item set that ended with [m] were combined with the phoneme [p] (frequency of word final coda cluster [mp]: 371 word types, based on the CELEX computerised database of Dutch) while the 11 words ending with [r] and the 5 ending with [ŋ] were combined with the phoneme [k] (frequency of word final coda cluster [rk]: 768 word types; frequency of word final coda cluster [ŋk]: 706 word types; e.g., [sɪste:mp] ’system+p’, [sɪχɑ:rk] ’cigar+k’ or [χɑŋk] ’floor+k’ respectively). For a full list of items see Appendix B.

In addition to the target-bearing experimental items, extra target-bearing fillers were also included in the lists in order to balance the number of mono-, bi-, and trisyllabic target-bearing strings. For the fricative item set 8 bisyllabic and 8 monosyllabic nouns were combined with nonsense syllables, yielding 16 extra items for the *syllabic* condition which would otherwise have been underrepresented within the experiment. For the *stop* item set 7 bisyllabic and 3 monosyllabic extra nouns were embedded in *syllabic* contexts. The construction of fillers in both item sets followed the same rationale as in Experiment 2.1A. For each set of items, six randomized lists were created. The counterbalancing of lists was the same as in Experiment 2.1A. Thus, each list contained all target words and all filler items, but the lists varied with respect to the contexts the target words appeared in. Thus, in the *stop* item set, each list contained 10 targets in each condition, yielding a total of 30 target-bearing experimental items (plus 10 extra target-bearing fillers for the syllabic condition). In the *fricative* item set there were 16 target-bearing strings per condition per list, yielding a total of 48 target-bearing experimental items in each list (plus 16 extra target-bearing items for the syllabic condition). For each list two randomizations were constructed that varied only in the order of stimuli presentation. The distance between target-bearing strings varied from one to three filler items.

Recording. Materials were recorded by the same female speaker who had already spoken the first sets of items. The recording was done in the same soundproofed booth with the same technical equipment that had been used for the first recording. The materials were measured and spliced into individual speech files, using the Xwaves/ESPS speech editor and were then transferred to a PC for use in the experiment.

Table 3.1:
Experiment 3.1: Mean word-spotting latencies (RT, ms) and Error Rates (%) in each of the three conditions in the stop item set.

	context		
	morphological	consonantal	syllabic
RT (ms)	676	673	699
Error rate (%)	7	16	10
Examples	kopiet/deurt	kopiek/deurp	kopiekel/deurtach

Subjects and Procedure. Thirty-six subjects participated in each version of the experiment (i.e., the *stop* and the *fricative* versions). In total, there were 59 women and 13 men, all of whom were students from Nijmegen University. They were all Dutch native speakers and they were all paid for their participation. Exactly the same experimental setup was used as in Experiment 2.1A. The only difference was that subjects had a short break after they had heard the first half of the stimuli. Each half of the experiment lasted about ten minutes.

Results

Stop items

Raw RTs measured from word onsets were adjusted by subtracting the lengths of the embedded target words in order to yield RTs from word offsets. As in the first experiment, responses where subjects either pressed the button and failed to give an answer at all or spotted an unintended word were treated as errors. This was true of 5.3% of the subjects' responses. Three items (*parool, teneur, geul*) were excluded from further analyses because of error rates higher than 50%. Table 3.1 and Figure 3.1 show the mean RTs and error rates per experimental condition.

The results of an Analysis of Variance (ANOVA) for RTs with participants (F_1) and items (F_2) as repeated measures did not show a significant main effect of context type but a highly significant effect of the factor number of syllables (i.e., whether the embedded word was mono- or bisyllabic; $F_1(1,29) = 77.75$, $p < .0001$; $F_2(1,25) = 31.47$, $p < .0001$). Only in the subject analysis was there a significant interaction of the factors syllable number and context type ($F_1(2,58) = 4.81$, $p < .05$; $F_2 < 1$). An ANOVA for the error data showed main effects of both context type ($F_1(2,66) = 4.14$, $p < .05$; $F_2(2,50) = 3.21$, $p < .05$) and syllable number ($F_1(1,33) = 39.37$, $p < .0001$; $F_2(1,25) = 19.83$, $p < .0001$) while there was no interaction of these factors.

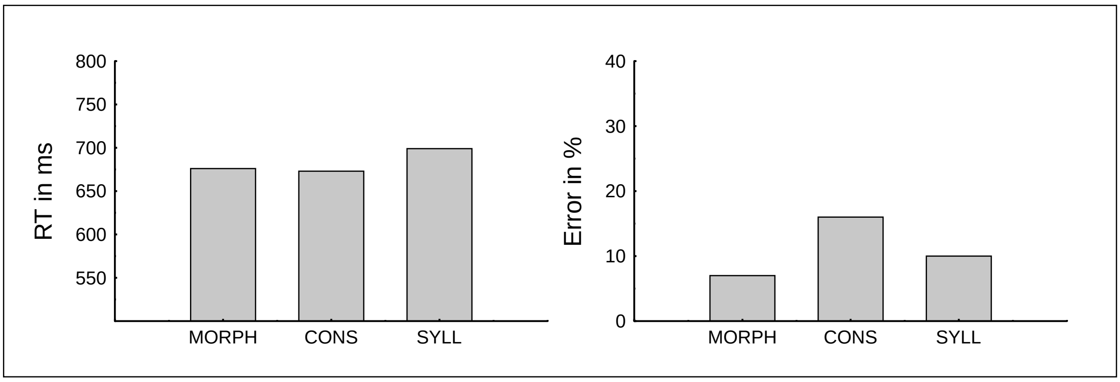


Figure 3.1: Experiment 3.1 - *Stop items*: Mean RTs (in ms) and mean error rates (in %) per experimental condition (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

An overall t-test on the error rates showed that the spotting of words in the *morphological* condition was significantly more accurate than the spotting of words in the *consonantal* condition ($t_1(35) = -3.36$, $p < .01$; $t_2(26) = -2.74$, $p < .01$) while no other pairwise comparison showed reliable differences.

As suggested by the highly significant factor syllable number, subjects were reliably faster in spotting bisyllabic words (mean RT: 604 ms) as compared to monosyllabic words (mean RT: 916 ms). Thus, bisyllabic targets were on average spotted 312 ms faster than monosyllabic targets. Also, the spotting of bisyllabic target words was much easier for subjects as they made on average only 10% errors as compared to 28% errors in the monosyllabic item set. This outcome is not surprising given the fact that bisyllabic words (mean length: 434 ms) were on average 180 ms longer than monosyllabic words (mean length: 254 ms). Because participants had significantly more processing time for the bisyllabic target words than for the shorter monosyllabic target words ($t(25) = -6.93$, $p < .0001$), they could initiate their decisions faster relative to the words' offsets. This interpretation is also supported by a reliable negative correlation between mean word length and RT (monosyllabic targets: $r(24) = -.66$, $p < .0001$; bisyllabic targets: $r(57) = -.44$, $p < .001$) implying that the longer an embedded word was, the shorter the reaction time became.

Individual ANOVAs were calculated separately for the mono- and bisyllabic item sets. Table 3.2 shows the respective RTs and error rates. There was no main effect of context type for either RTs or error rates within the monosyllabic item set. Note that this set of items was particularly small, with only 8 items contributing to the data (because the item *geul* was excluded) and thus low power probably prevented any potential effects from reaching significance.

Table 3.2:
Experiment 3.1: Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions for the stop item set listed separately for mono- and bisyllabic target words.

monosyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	951	922	869
Error rate (%)	17	27	26
Example	deurt	deurp	deurtach
bisyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	578	587	645
Error rate (%)	2	11	3
Example	kopiet	kopiek	kopiekel

For the bisyllabic item set there was a main effect of context type for error rates ($F_1(2,66) = 11.56$, $p < .0001$; $F_2(2,36) = 4.06$, $p < .05$) but no such effect for the RTs. A t-test showed that participants made significantly more errors in the *consonantal* condition than in the *morphological* condition ($t_1(35) = -3.27$, $p < .01$; $t_2(18) = -2.14$, $p < .05$). The difference between the *consonantal* and the *syllabic* condition was nearly significant ($t_1(35) = 3.12$, $p < .01$; $t_2(18) = 2.05$, $p < .06$) with the latter being more accurate than the former. There was no difference between the *morphological* and the *syllabic* conditions.

Fricative items

Adjusted RTs and error rates for the *fricative* item set are listed in Table 3.3 by experimental condition and are also shown in Figure 3.2. Again, responses where subjects either pressed the button and failed to give an answer at all or spotted an unintended word were treated as errors. This was true for 2% of the subjects' reactions. Five items (*parool, teneur, bil, gong, tang*) had to be excluded from further analyses because they produced error rates of more than 50%.

An overall ANOVA for RTs showed a main effect of context type for both subjects and items as repeated measures (F_1 $(2,68) = 7.73$, $p < .0001$; $F_2(2,82) = 4.62$, $p < .001$). The factor syllable number was also highly significant ($F_1(1,34) = 167.51$, $p < .0001$; $F_2(1,41) = 45.44$, $p < .0001$). No reliable interaction of these

Table 3.3:
Experiment 3.1: Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions for the fricative item set.

	context		
	morphological	consonantal	syllabic
RT (ms)	649	638	713
Error rate (%)	12	15	12
Examples	sigaars/duims	sigaark/duimp	sigaarkief/duimfoel

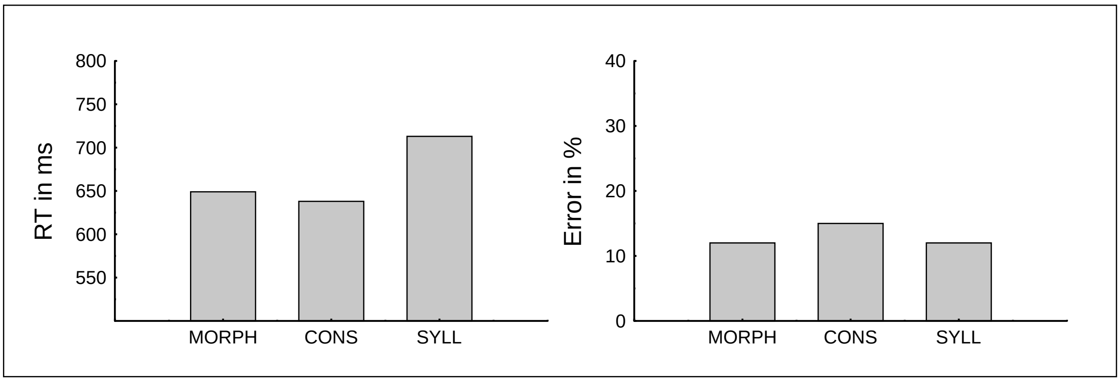


Figure 3.2: Experiment 3.1 - *Fricative items*: Mean RTs (in ms) and mean error rates (in %) per experimental condition (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

factors was observed. The results of an overall t-test showed that mean RTs in the *syllabic* condition were significantly slower than both in the *consonantal* (t_1 (36) = -2.5, $p < .05$; $t_2(42) = -2.52$, $p < .05$) and the *morphological* condition ($t_1(36) = -2.5$, $p < .05$; $t_2(42) = -2.47$, $p < .05$). There was no significant difference between the RTs in the latter two conditions. In the error analysis only the factor number of syllables showed significant effects ($F_1(1,34) = 36.99$, $p < .0001$; $F_2(1,41) = 17.23$, $p < .0001$). There was neither a main effect of context type nor an interaction of the two main factors.

As in the previous experiment, bisyllabic target words were spotted reliably faster and more accurately than monosyllabic target words. On average, participants needed only 549 ms to spot bisyllabic words and were thus 254 ms faster than they were in spotting monosyllabic words (mean RT: 813 ms). Further, bisyllabic targets were spotted more accurately (mean error rate: 10 %) than monosyllabic targets (mean error rate: 24 %). The same explanation as for the *stop* item set can be given: the longer bisyllabic words (mean length: 421 ms) gave listeners significantly more processing time (t (41) = -8.51, $p < .0001$) as

Table 3.4:
Experiment 3.1: Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions for the fricative item set listed separately for mono- and bisyllabic target words.

monosyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	806	769	862
Error rate (%)	15	24	21
Example	duims	duimp	duimfoel
bisyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	510	537	598
Error rate (%)	8	7	3
Example	sigaars	sigaark	sigaarkief

compared to the relatively short monosyllabic words (277 ms) and thus enabled them to give faster responses relative to the words' offsets. Again, the data set was analysed separately for mono- and bisyllabic target words (Table 3.4) and individual ANOVAs were calculated for both sets of items.

The individual ANOVA for monosyllabic target words did not show any significant main effects, neither for RTs nor for error rates. For the bisyllabic item set, however, there was an almost significant effect of the factor context type for RTs ($F_1(2,68) = 13.83$, $p < .0001$; $F_2(2,42) = 2.93$, $p < .06$). No such effect was observed for the error rates. A t-test showed that bisyllabic words embedded in *consonantal* contexts were spotted significantly faster than the same words embedded in *syllabic* contexts ($t_1(36) = -3.15$, $p < .01$; $t_2(21) = -2.57$, $p < .05$). The same was true for the pairwise comparison of the *morphological* with the *syllabic* condition although only the subjects analysis was statistically fully reliable (t_1 $(36) = -4.40$, $p < .0001$; $t_2(21) = -1.85$, $p < .08$).

Because the results for *stop* and *fricative* items had been different in Experiment 2.1A, the factor 'item set' was included in an overall ANOVA calculated for the combined data of *stop* and *fricative* item sets of the current experiment. This factor was not significant ($F_1 < 1$) and also the interaction with the factor context type was not significant ($F_1 = 1$).

Discussion

Overall, the pattern of results looks similar for both item sets. The difference between *stop* and *fricative* items that was indicated in Experiment 2.1 was not replicated here. This supports the explanation that was given for the missing PWC effect in the *fricative* item set of Experiment 2.1. Remember that subjects were not slower or less accurate when they had to spot words that were embedded in *consonantal* contexts in the fricative item set than when they had to spot these words embedded in *syllabic* contexts. A different pattern was observed for the *stop* item set of Experiment 2.1, where there was a difference between these conditions in the predicted direction: words embedded in *consonantal* contexts were spotted less accurately and slower than the same words embedded in *syllabic* contexts.

It was therefore hypothesized that participants in Experiment 2.1 were sensitive to the infrequent consonant clusters *rf* and *mf* at the codas of the strings (e.g., *deurf* or *duimf*) and could therefore solve the task on the basis of phonotactic information alone. These consonant clusters were thus exchanged for more frequent ones in the latest experiment in order to deal with this phonotactic confound. The similar patterns for the *stop* and *fricative* item sets in the current experiment indicate that the difference observed for these item sets in Experiment 2.1 was indeed due to a phonotactic confound rather than to a general processing difference between the two sets.

This does not explain, however, why the expected PWC effect was reversed in the current experiment. That is, in both experiments participants were slowest when they spotted words embedded in a possible word context, i.e. the context that is supposed to make the spotting of words easiest and thus fastest. The only exception to that result was observed for the monosyllabic nouns in the *stop* item set. Here, words embedded in a possible word context were spotted faster than the same words in the other contexts. This result, however, was not statistically reliable. Overall, as illustrated in Figures 3.1 and 3.2, the RT data in the *morphological* condition clusters with that in the *consonantal* condition. Disregarding the lack of the basic PWC effect for the moment, this could be interpreted as a hint that morphemes are not treated differently from non-morphological consonants. This conclusion, however, is contradicted by the pattern of the error rates. As predicted by the PWC, words embedded in possible (i.e., *syllabic*) contexts are spotted more accurately than words in impossible (i.e., *consonantal*) contexts. But words in *morphological* contexts were also spotted more accurately than words in *consonantal* contexts.

The error data of the *morphological* condition therefore suggests exactly the opposite of what the respective RT data implies: namely, that morphemes have a special status in the segmentation process. It is thus easier for listeners to spot words when the residue of the segmentation process is a morpheme of that language as compared to a ’meaningless’ consonant. This contradiction between the RT and error data makes it impossible to draw strong conclusions about the status of morphemes in the segmentation process. Furthermore, there is a potential speed-accuracy trade-off in the results in the *syllabic* condition: although RTs were slowest here, responses were the most accurate. The reversal of the predicted effect is therefore present in the RT data but not in the error data.

In the light of these contradictory results, two major questions arise: first, why are words embedded in the condition that was predicted to be the easiest spotted slowest, and second, why are those words on the other hand spotted most accurately? A closer inspection of mono- and bisyllabic target words could probably supply a solution to at least the first of the two questions. The pattern of RTs in both *stop* and *fricative* item sets seems mainly to be due to the bisyllabic targets. The overall pattern of RTs in the *stop* item set is clearly only attributable to the bisyllabic words as for the monosyllabic words the RTs in the *syllabic* condition are faster than in the other two conditions (even if not significantly so). In the *fricative* item set, although RTs in the monosyllabic item set are slowest in the *syllabic* condition, this result is again not supported by the statistics. In the bisyllabic *fricative* item set, however, the *syllabic* condition is significantly slower than both the *consonantal* and the *morphological* conditions. These results suggest that the bisyllabic items are mainly responsible for the pattern of the RT data in both experiments.

Careful auditory inspection of the bisyllabic items in the different conditions revealed that the speaker had stressed the second (strong) syllables of bisyllabic embedded target words more emphatically when she had produced them in *syllabic* contexts (e.g., *sigaar* in *sigaarkief*) than when she had uttered them in either of the other two contexts (e.g., *sigaar* in *sigaars, sigaark*). Remember that Dutch is a stress-timed language which means that listeners tend to assume word boundaries before strong syllables. It is thus possible that listeners in the *syllabic* condition tended to assume a word onset before the strong second syllable as well as at the actual word onset. For example, the second strong syllable of the word *mobiel* (i.e., [bi:l]) could temporarily activate words like *biologie, bison, bilateral* that would temporarily compete for recognition with the intended word. Because of this competition, recognition of the correct words would be inhibited. This ’competition effect’ would be less effective in the other two conditions because the second syllables of the embedded words were less strongly stressed as compared to the *syllabic* condition. The reduction of the activation of the actually intended words in the *syllabic* contexts may thus have resulted in the longer RTs. The influence of metrical information on the task could thus have masked the influence of the PWC in the *syllabic* condition.

In order to test whether some acoustic confound masked the PWC effect, Experiment 3.1 was run again, but with cross-spliced versions of all items. Thus, the same version of each target word was used in the three experimental conditions. For example, the word *sigaar* was spliced out of one natural waveform and then cross-spliced with the three different contexts (e.g., *-t, -k, -kief*). Any acoustic factors within the embedded words were therefore kept constant across all conditions. Any potential influence of these acoustic factors on the task should thus be effective in all contexts.

Experiment 3.2: Word-Spotting - Control Experiment

Method

Materials and Recording. Exactly the same materials as in the previous experiment were recorded again by the same female speaker and were cross-spliced afterwards using the Xwaves /ESPS speech signal processing package. Recording facilities were the same as in Experiment 3.1. The speaker produced each target word in each condition several times so that there was a broad selection from which the most natural-sounding version of each target word could be chosen. This choice was based on careful auditory inspection. The selected target word was then excised from the context it was originally produced in and cross-spliced afterwards with the three different contexts that were all excised from other utterances than the embedded word. None of the targets was spliced out of a *syllabic* context in order to avoid the potential prosodic confound that might have been responsible for the pattern of RTs in Experiment 3.1. Half of the items in each item set (i.e., 24 in the *fricative* and 15 in the *stop* item set) were excised from a *consonantal* context while the other half were excised from a *morphological context*. Thus, there were no naturally uttered items in the experiment: all were cross-spliced. All cuts were made at or close to zero crossings with the constraint that this would result in smooth transitions between the words' offsets and the following contexts' onsets. None of the splices could be heard.

Procedure and Subjects. The newly created items were presented to listeners in exactly the same randomized orders with exactly the same experimental setup as before. 24 - 18 female and 6 male - students from Nijmegen University were paid for their participation. Half of them were tested on the *stop* item set while the other half were presented with the *fricative* item set.

Table 3.5:
Experiment 3.2 (control experiment): Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions for the stop item set with cross spliced items.

	context		
	morphological	consonantal	syllabic
RT (ms)	578	594	671
Error rate (%)	21	18	15
Examples	kopiet/deurt	kopiek/deurp	kopiekel/deurtach

Results

Stop items

As can be seen in Table 3.5, where the adjusted RTs and error rates for each condition are listed, the same pattern of results was found as in Experiment 3.1. Responses where subjects pressed the button but spotted the wrong word or failed to give a vocal response at all were counted as errors. A proportion of 2.7% of all trials was rejected for that reason. Only one item (*parool*) was excluded from further analyses, because of a total error rate of 75%. Four other items with error rates between 50% and 60% (*kopie, tor, wang, geul*) were not excluded because of the already limited number of data points. Although the data set was much smaller than in the former experiment (only 12 subjects) there was a strong main effect of the factor syllable number in an ANOVA for both RTs and error rates (error rates: $F_1(1,9) = 12.94$, $p < .01$; $F_2(1,27) = 7.91$, $p < .01$; RTs: $F_1(1,9) = 46.11$, $p < .0001$; $F_2(1,27) = 20.87$, $p < .0001$). No main effect of context type was observed for either the RT data or for the error data and there was no interaction of the two main factors. But especially the RT data showed exactly the same result pattern as was observed in Experiment 3.1: RTs in the *syllabic* condition were again slower than in the other two conditions. There was a significant difference in mean word length between monosyllabic target words (305 ms) and bisyllabic target words (471 ms) explaining the significant effect of the factor syllable number on both RTs and error rates ($t(27) = -6.56$, $p < .0001$). Again - as is listed in Table 3.6 - the data was split up by mono- and bisyllabic items. In ANOVAs performed separately for both mono- and bisyllabic items the factor context type was not significant for either of the dependent variables.

Table 3.6:
Experiment 3.2 (control experiment): Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions for the stop item set with cross spliced items and listed separately for mono- and bisyllabic target words.

monosyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	769	785	827
Error rate (%)	28	28	31
Example	deurt	deurp	deurtach
bisyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	507	525	618
Error rate (%)	18	14	8
Example	kopiet	kopiek	kopiekel

Fricative items

Table 3.7 shows adjusted mean RTs and error rates per experimental condition for the *fricative* item set. A proportion of 2.1% of all trials was rejected as errors because subjects pressed the button but either failed to give a correct response or gave no response at all. Two items (*parool, gong*) were excluded from further analyses because of error rates higher than 60%. One other item (*vier*) with an error rate of 50% was however not excluded, in order to keep the data set as large as possible.

For the RT analysis there was an almost significant effect of context type ($F_1(2,18) = 7.62$, $p < .01$; $F_2(2,88) = 2.91$, $p < .06$) and a strong main effect of the factor syllable number (F_1 $(1,9) = 33.56$, $p < .0001$; $F_2(1,44) = 44.39$, $p < .0001$). An ANOVA for error rates only revealed a significant main effect of the factor syllable number ($F_1(1,9) = 13.06$, $p < .01$; F_2 $(1,44) = 21.06$, $p < .0001$). A t-test performed on the RT data showed that items embedded in *morphological* context were spotted significantly faster than the same items embedded in *syllabic* context ($t_1(11) = -6.31$, $p < .0001$; $t_2(45) = -2.60$, $p < .01$). No other pairwise comparison was significant.

Target words in the bisyllabic item set (mean length: 456 ms) were significantly longer than target words in the monosyllabic item set (mean length: 303 ms; $t(44) = -8.82$, $p < .0001$). The item set was again analysed separately for mono- and bisyllabic targets as listed in Table 3.8. ANOVAs performed separately for each

Table 3.7:
Experiment 3.2 (control experiment): Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions for the fricative item set with cross-spliced items.

	context		
	morphological	consonantal	syllabic
RT (ms)	556	586	659
Error rate (%)	15	20	15
Examples	sigaars/duims	sigaark/duimp	sigaarkief/duimfoel

group of items only showed a significant main effect of context type for the bisyllabic item set for RTs ($F_1(2,18) = 7.40$, $p < .01$; $F_2(2,44) = 5.01$, $p < .01$) but not for error rates. No reliable effects for the monosyllabic item set were obtained. A t-test on RTs in the bisyllabic item set showed that RTs in the *morphological* condition and in the *consonantal* condition, respectively, were significantly shorter than in the *syllabic* condition, but were not significantly different from each other (*morphological* vs. *syllabic*: $t_1(11) = -4.93$, $p < .0001$; $t_2(22) = -3.28$, $p < .01$; *consonantal* vs. *syllabic*: $t_1(11) = -3.23$, $p < .01$; $t_2(22) = -2.39$, $p < .05$).

Discussion

The results of Experiment 3.2 convincingly show that the prosodic confound in the target words in Experiment 3.1 did not cause the reversed PWC effect. Remember that the PWC predicts that words embedded in *syllabic* contexts should elicit the fastest RTs and the lowest error rates as compared to the same words embedded in *consonantal* contexts. The opposite result was found in Experiment 3.1. Even if not all the results in Experiment 3.2 were statistically reliable (because of the small number of data points), they showed the same pattern as those in Experiment 3.1 with naturally uttered stimuli. The condition that was expected to produce the fastest RTs again proved to be the slowest condition.

If the PWC effect was not masked by a metrical confound, the question still has to be answered why listeners in these experiments produced slower RTs in a condition the PWC theory predicts they should be fastest in. Three factors - that are somewhat interdependent - might be responsible for this outcome. One factor is connected to the relative early uniqueness points of the bisyllabic target words. Another is related to the position of target word embeddings. The third factor has to do with the different context lengths in the three conditions.

Table 3.8:
Experiment 3.2 (control experiment): Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions for the fricative item set with cross-spliced items and listed separately for mono- and bisyllabic target words.

monosyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	680	727	765
Error rate (%)	26	23	23
Example	duims	duimp	duimfoel
bisyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	461	464	574
Error rate (%)	4	11	3
Example	sigaars	sigaark	sigaarkief

Early uniqueness

All bisyllabic words used in the current study were unique at the penultimate phoneme at the latest (with one exception, namely the target *forel* ’trout’, which is unique only at its last phoneme [l]). On average, the bisyllabic targets were 5 phonemes long and were unique at the fourth phoneme. The earlier a certain word reaches its uniqueness point, the earlier other cohort competitors (i.e., lexical candidates beginning at the same point as the targets) will be excluded from the shortlist of competitors. And as soon as the last of these competitors has been dropped from the shortlist, identification of the intended target is possible. Thus, it is likely that listeners in the last two experiments were able to identify the target words on most trials before they had even heard the last phoneme of those target words.

The consequence for the word-spotting task is quite obvious: this task can only expose an effect of the PWC with following context if the identification of an embedded word is delayed until this very context is processed. Only if the word cannot be identified before the following context is encountered can this following context modulate the segmentation process. Imagine that, on average, listeners could indeed identify the embedded words before or right at the last phoneme of the target word, irrespective of the following context it was embedded in. The target word *probleem* ’problem’, for example, is already unique at the phoneme [b] and can thus be identified three phonemes before the context that is supposed to constrain identification via segmentation is actually taken into account. Lexical identification can thus take place before the experimental manipulation can start influencing the identification of the word. Strictly speaking, *segmentation* was no longer necessary for listeners to be able to perform the task successfully. Because words could be identified early, listeners did not need to segment them out of the context they were embedded in. Following context of any kind was not effective because the word had already been identified. In some sense, it was too easy for subjects to identify the target words, which in turn prevented the PWC from exerting an influence on listeners' performance. This was strengthened by another facilitatory factor: the constant position of target embeddings.

Initial embeddings

Listeners in the current experiments could rely on the fact that the onsets of target words always matched the onsets of the nonsense strings they were embedded in. Thus, if a stimulus contained an embedded word this word always started at the first phoneme of that stimulus. The task was thus simplified by this factor. The question then arises why there were RT and error rate differences between the conditions at all if the task was so easy to solve. Intuitively, one would expect equal results in all conditions if following context was ignored. As we saw, however, the actual results tell a different story.

Context length

The unexpectedly longer RTs in the *syllabic* condition suggest that the factor context had an unintended effect on listeners' performance. The factor context varied not only with respect to the feature '(im)possible word' but also with respect to the feature 'length': syllables are by nature longer than single consonants. The mean context-lengths in each condition are listed in Table 3.9 for the two item sets in Experiments 3.1 and 3.2. Thus, the results that were obtained in the latest experiments seem to vary as a function of the feature 'length' rather than the feature 'possible word'. Positive correlations of RTs with the following context-lengths in some conditions support this interpretation (**Experiment 3.1**: *fricative* item set/*syllabic* condition: r(43) = .39, $p < .01$; *stop* item set/*morphological* condition: r(27) = .40, $p < .05$; *syllabic* condition: r(27) = .37, $p < .06$; **Experiment 3.2**: *fricative* item set/*syllabic* condition: r(46) = .35, $p < .01$). The longer the following context was, the longer the RTs became. What was measured in these experiments was not the ease of the segmentation but rather the time listeners waited before they initiated their responses: the longer the context was, the longer subjects delayed their reactions. Hence longer RTs were

Table 3.9:
Experiments 3.1 and 3.2: Mean context lengths (in ms) in each of the three conditions across item sets.

stop item set			
		context	
	morphological	consonantal	syllabic
Mean context length (ms)			
Experiment 3.1	205	218	455
Experiment 3.2	212	207	474

fricative item set			
		context	
	morphological	consonantal	syllabic
Mean context length (ms)			
Experiment 3.1	329	264	435
Experiment 3.2	320	265	459

obtained in a condition that was predicted to be the fastest.

ANCOVAs with the factor ’context length’ as covariate were performed separately for each set of items (*stop* and *fricative* item sets) in each experiment (3.1 and 3.2). These had to be calculated for mono- and bisyllabic target words separately because the number of items varied between mono- and bisyllabic targets. All context effects on RTs that had been observed in the ANOVAs (i.e., **Experiments 3.1 and 3.2**: *fricative* item set/bisyllabic targets) vanished when the factor ’context-length’ was included as a covariate (**Experiment 3.1**/*fricative* item set/bisyllabic targets: $F_2 < 1$; **Experiment 3.2**/*fricative* item set/bisyllabic targets: $F_2 < 1$).

But how does this interpretation correspond with the earlier argument that listeners were able to identify embedded target words relatively early? If this were indeed the case, why should they wait until the end of each string before they pressed the response button? There is only a speculative answer to that question. It seems that the ability to identify the embedded words early gave subjects in some sense the *luxury* to wait until they had heard the whole string. One reason for this waiting-strategy might be that they used the extra time to double-check whether they had really spotted the right word. They may have wanted to be sure that no even longer word was embedded. This explanation is not unlikely given the high rate of compounding that occurs in Dutch. Note that this explanation can only hold if the hypothesis is correct that listeners could identify the embedded target words early (i.e., before the following context was encountered). If they had not recognised the words early, they would have been confronted with a segmentation problem to solve. If so, differences between the three experimental conditions in the predicted directions should have been observed.

The next two experiments were designed to control for these confounds. In Experiment 3.3 I tried to make the task more demanding, in order to delay identification so that the following contexts would be taken into account. Experiment 3.4, on the other hand, attempted to take care of the length problem.

Experiment 3.3: Spotting words embedded both item-initially and item-finally

In order to impede performance in the word spotting task, both initial and final embeddings (e.g., *lepel* 'spoon' in *blepel* or *kulepel*) were presented in Experiment 3.3. So far, words in Experiments 3.1 and 3.2 had only been embedded item-initially, and this might have led subjects to an attentional strategy: they could concentrate on the onsets of nonsense strings to look specifically for words to begin at those onsets. A variable target position blocks this sort of strategy. Participants could thus not longer rely on the fact that the onsets of embedded words always matched the onsets of the nonsense strings as they could do in Experiments 3.1 and 3.2. Apart from this change, the predictions of the PWC were the same as in the earlier experiments: *syllabic* contexts (either following or preceding the target word) should make the location of word-boundaries easy. Therefore, the spotting of the word *lepel*, for example, should be faster and more accurate in a nonsense string like *kulepel* than in a string like *blepel*. The same pattern of results should be obtained for items that are embedded at the initial position of nonsense strings (i.e., *probleem* in *probleemp* and *probleemtaaf*). The logic of this new set-up was simply that delayed identification, as the result of a more difficult task, might slow subjects' responses down and therefore allow following context to influence word-spotting performance.

Method

Materials. For each item set, the number of nonsense strings that had words embedded item-initially and those that had item-final embeddings was balanced. There were thus 30 item-initial and 30 item-final embeddings in the *stop* item set, while there were 48 items with final embeddings and 48 strings with initial embeddings in the *fricative* item set.

For initial embeddings, exactly the same *stop* and *fricative* item sets were used as in the previous experiments. For final embeddings, a new set of bisyllabic words was included in the item lists. In order to balance the number of nonsense strings that had words embedded at initial and at final position, 48 new items were included in the *fricative* item set. A subset of those 48 words - 30 in total - was included in the *stop* item-list. Corresponding to the items with initial embeddings, there was a possible word condition with a full syllable preceding the target words (e.g., *loper* 'runner' in *zoeloper*) as well as an impossible word condition with only a single consonant as preceding context (e.g., *loper* in *bloper*). Because in Dutch there are no single consonant morphemes at the beginnings of words, the third condition was a reduced syllable (Cə) like *loper* in *keloper*. The main reason for using three different contexts for final embedded words was to keep them as comparable as possible to the initial embedded items that also occurred in three different kinds of contexts. Final embedded target words had either the stress pattern SW (86 % of the items) or WS (14 % of the items). For a full list of new items with final embeddings see Appendix B.

The proportion of fillers within the different item sets also had to be balanced. New fillers were constructed to match structurally the three different nonsense strings with final embeddings. For the *fricative* item set, 96 new filler items were included: one third was bisyllabic while the rest consisted of three syllables. 60 of those were included in the *stop* item set, again with one third being bisyllabic and the rest trisyllabic. None of the fillers had real Dutch words embedded within them. Per item set, three experimental lists were constructed that contained all target words and filler items. The lists differed with respect to the contexts the targets were embedded in. In the *fricative* item set, each of the six experimental conditions (e.g., three with following and three with preceding context) contained 16 target-bearing strings, while there were 10 target-bearing strings per condition in the *stop* item set. Two randomizations of each list were created so that the order of presentation was different between the lists. There were thus six different experimental lists per item set. The randomizations were constrained by one factor: there was at least one filler item between two target bearing strings and there were never more than three fillers between two experimental items. The same female speaker as in previous experiments recorded the materials. Recording facilities and location were identical to those in Experiments 3.1 and 3.2.

Subjects and Procedure. 72 students from Nijmegen University were paid for their participation in one of the two experiments. Half of them were presented with the *fricative* item set while the other half were tested on the *stop* item set. The experimental set-up was exactly as in the previous experiments. The only

Table 3.10:
Experiment 3.3: Mean word-spotting latencies (RT, in ms), Error Rates (%), and mean context lengths (in ms) in each of the three conditions for initially embedded words in the stop item set.

	context		
	morphological	consonantal	syllabic
RT (ms)	596	560	659
Error rate (%)	11	11	20
Mean context length (ms)	217	179	396
Examples	kopiet/deurt	kopiek/deurp	kopiekel/deurtach

difference was that the experiment took longer in total due to the extended set of items. There was a brief practice phase followed by the main session, which was divided into three parts, giving subjects the opportunity to pause two times in total. Each session took about 50 minutes.

Results

Initial embeddings: *Stop* items

In Table 3.10 adjusted mean RTs and error rates are listed for each condition in the *stop* item set. These means are also plotted in Figure 3.3. As in previous experiments, responses where subjects pressed the button but either failed to give an answer or spotted the wrong word were counted as errors (1.6%). Four items (*teneur, tor, peul, bil*) were excluded from further analyses because of error rates higher than 50%. In an overall ANOVA the factor *syllable number* was not taken into account because half of the subjects (18 in total) failed to spot any monosyllabic target words in at least one of the three conditions. For the same reason an individual ANOVA was only performed on the bisyllabic item set. I will come back later to a comparison between the overall ANOVA and the individual ANOVA for bisyllabic targets in order to discuss the influence of monosyllabic target words on the overall pattern.

In the overall ANOVA there was a highly significant main effect of context type for both RTs and error rates (RTs: $F_1(2,66) = 9.72$, $p < .0001$; $F_2(2,50) = 9.78$, $p < .0001$; Error Rates: $F_1(2,66) = 9.12$, $p < .0001$; $F_2(2,50) = 3.29$, $p < .05$). Pairwise comparisons of the three conditions showed that both the *morphological* and the *consonantal* condition produced reliably faster RTs as compared to the *syllabic* condition (*morphological* vs. *syllabic*: $t_1(35) = -2.63$, $p < .01$; $t_2(25) = -2.94$, $p < .01$; *consonantal* vs. *syllabic*: $t_1(35) = -4.86$, $p < .0001$; $t_2(25) = -3.96$,

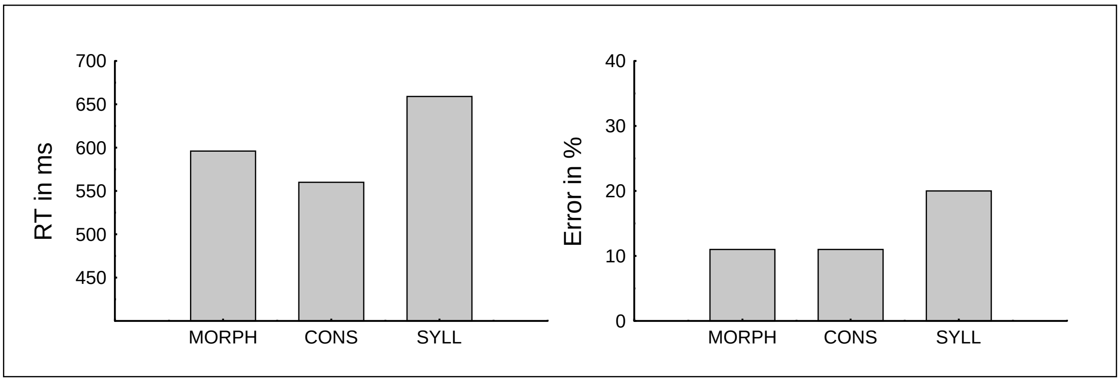


Figure 3.3: Experiment 3.3 - *Initial embeddings/Stop items*: Mean RTs (in ms) and mean error rates (in %) per experimental condition (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

$p < .01$). Despite the 36 ms RT difference between the *morphological* and the *consonantal* conditions, the pairwise comparison of those conditions was not significant.

Almost the same pattern was observed for the error data: participants made significantly fewer errors in the *consonantal* as compared to the *syllabic* condition ($t_1(35) = -3.22$, $p < .01$; $t_2(25) = -2.33$, $p < .05$) while the same tendency was not fully reliable for the comparison between the *morphological* and the *syllabic* conditions ($t_1(35) = -3.15$, $p < .01$; $t_2(25) = -1.86$, $p < .1$). Again, there was no difference whatsoever between the *morphological* and the *consonantal* condition. Although the factor syllable number was not taken into account in these analyses, the data was analysed separately for mono- and bisyllabic targets, as is listed in Table 3.11.

As already observed in the previous word-spotting experiments, bisyllabic targets (mean length: 530 ms) were significantly longer than monosyllabic target words (mean length: 395 ms; $t\ (24) = -5.32$, $p < .0001$). The ANOVA calculated individually for the bisyllabic item set showed a main effect of context type for RTs ($F_1(2,66) = 11.12$, $p < .0001$; $F_2(2,38) = 8.88$, $p < .001$). In the error analysis, however, there was no significant effect of the factor context type. The result of a t-test showed exactly the same pattern of results for RTs as the overall pairwise comparisons: both the *morphological* and the *consonantal* condition produced reliably faster RTs than the *syllabic* condition (morphological vs. *syllabic*: $t_1(35) = -2.89$, $p < .01$; $t_2(19) = -2.80$, $p < .01$; *consonantal* vs. *syllabic*: $t_1(35) = -5.05$, $p < .0001$; $t_2(19) = -3.70$, $p < .01$). No difference in RTs was observed for the *morphological* condition as compared to the *consonantal* condition.

Because the RT pattern does not change when the monosyllabic items are excluded, one can assume that both item sets behave similarly in terms of RT. The overall error pattern, however, seems to be strongly influenced by the monosyl-

Table 3.11:
Experiment 3.3: Mean word-spotting latencies (RT, in ms), Error Rates (%), and mean context lengths (in ms) in each of the three conditions for the stop item set with initially embedded targets listed separately for mono- and bisyllabic target words.

monosyllabic targets		context	
	morphological	consonantal	syllabic
RT (ms)	753	762	901
Error rate (%)	18	26	53
Mean context length (ms)	226	195	360
Example	deurt	deurp	deurtach

bisyllabic targets		context	
	morphological	consonantal	syllabic
RT (ms)	554	513	621
Error rate (%)	9	6	10
Mean context length (ms)	208	175	402
Example	kopiet	kopiek	kopiekel

labic item set, because the factor context type was highly significant in the overall analysis but not in the individual bisyllabic analysis. As is shown in Table 3.11 the error rates for monosyllabic target words in the *syllabic* condition are extremely high (even though all target words that produced overall error rates of more than 50% were excluded from the analysis). There were positive correlations of the factor context length with RTs in the *morphological* (r(26) = .54, $p < .01$) and the *consonantal* (r(26) = .44, $p < .05$) conditions. An ANCOVA with context-length as a covariate calculated separately for the bisyllabic item set produced no significant effect of the factor context type on RTs (F_2 = 2.07, n.s.). Thus, when context length was taken into account, the effect of context type on RTs vanished.

Initial embeddings: *Fricative* items

The adjusted RTs and error rates for the *fricative* item set are listed in Table 3.12 by experimental condition and plotted in Figure 3.4. Those trials where participants pressed the button but failed to give a correct vocal response were treated as errors (2.4%). Five items (*teneur, nul,wang, bil, kier*) were excluded from any further analyses because they produced error rates of more than 50%.

The result of an overall ANOVA showed reliable context effects for both RTs and error rates (RTs: $F_1(2,42) = 9.76$, $p < .0001$; $F_2(2,82) = 15.27$, $p < .0001$;

Table 3.12:
Experiment 3.3: Mean word-spotting latencies (RT, in ms), Error Rates (%), and mean context lengths (in ms) in each of the three conditions for initially embedded words in the fricative item set.

	context		
	morphological	consonantal	syllabic
RT (ms)	739	747	834
Error rate (%)	11	13	20
Mean context length (ms)	389	305	465
Examples	sigaars/duims	sigaark/duimp	sigaarkief/duimfoel

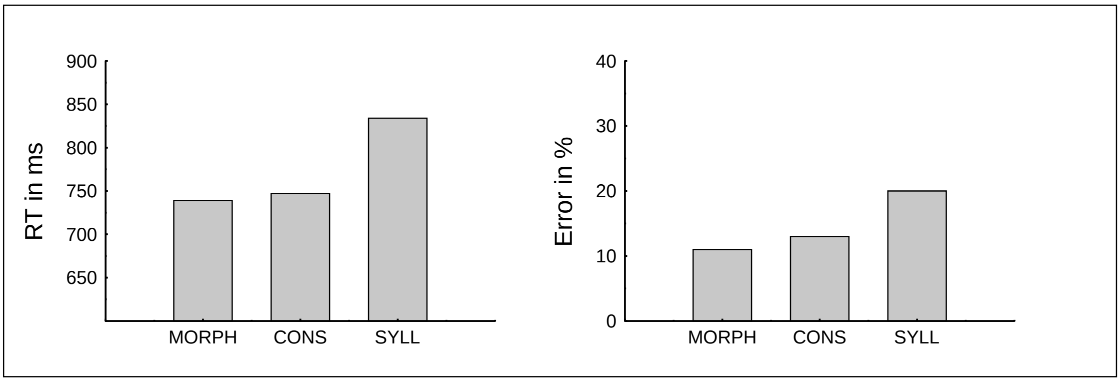


Figure 3.4: Experiment 3.3 - *Initial embeddings/Fricative items*: Mean RTs (in ms) and mean error rates (in %) per experimental condition (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

error rates: $F_1(2,42) = 5.61$, $p < .01$; $F_2(2,82) = 5.77$, $p < .01$). Also the factor number of syllables was highly significant in both analyses (RTs: $F_1(1,21) = 174.24$, $p < .0001$; $F_2(1,41) = 44.21$, $p < .0001$; error rates: $F_1(1,21) = 112.15$, $p < .0001$; $F_2(1,41) = 40.19$, $p < .0001$). There was no interaction of these factors in either the RT- or the error-analyses.

In an overall t-test on RTs, the *syllabic* condition was again significantly slower than both the *morphological* and the *consonantal* conditions while the latter two did not differ from each other (*morphological* vs. *syllabic*: $t_1(23) = -3.74$, $p < .01$; $t_2(42) = -4.70$, $p < .0001$; *consonantal* vs. *syllabic*: t_1 $(23) = -2.40$, $p < .05$; $t_2(42) = -3.93$, $p < .0001$). In terms of accuracy, the *syllabic* condition differed significantly from the other two conditions while there was again no difference in error rates between the *morphological* and the *consonantal* conditions. Error rates in the *syllabic* condition were significantly higher than in the *morphological* condition ($t_1(23) = -3.24$, $p < .01$; $t_2(42) = -3.00$, $p < .001$), while the same tendency was not fully reliable in the *syllabic-consonantal* comparison ($t_1(23) =$

Table 3.13:
Experiment 3.3: Mean word-spotting latencies (RT, in ms), Error Rates (%), and mean context lengths (in ms) in each of the three conditions for the fricative item set with initially embedded targets listed separately for mono- and bisyllabic target words.

monosyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	897	890	1045
Error rate (%)	20	23	36
Mean context length (ms)	388	258	459
Example	duims	duimp	duimfoel

bisyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	627	646	709
Error rate (%)	3	5	7
Mean context length (ms)	389	338	468
Example	sigaars	sigaark	sigaarkief

-2.22, $p < .05$; $t_2(42) = -1.93$, $p < .1$). The data was again analysed separately for mono- and bisyllabic targets. Table 3.13 shows the respective RT and error data.

As suggested by the highly significant main effect of the factor syllable number, bisyllabic targets were spotted faster and more accurately than monosyllabic targets. This was again due to the fact that bisyllabic targets (mean length: 478 ms) were on average significantly longer than monosyllabic target words (mean length: 325 ms; $t\,(41) = -7.97$, $p < .0001$) which gave participants significantly more processing time.

Monosyllabic target words

The pattern of results observed for the overall data set was also found for the monosyllabic item set alone. An ANOVA showed a reliable effect of context type for both RTs and error rates (RTs: $F_1(2,42) = 5.00$, $p < .01$; $F_2(2,38) = 8.35$, $p < .001$; error rates: $F_1(2,42) = 4.28$, $p < .05$; $F_2(2,38) = 4.00$, $p < .05$). Pairwise comparisons of the three conditions for only the monosyllabic target words showed that RTs in the *morphological* condition were significantly shorter as compared to the *syllabic* condition ($t_1(23) = -2.50$, $p < .05$; $t_2(19) = -3.59$, $p < .01$). While the *consonantal* condition was also significantly faster than the *syl-*

labic condition ($t_1(23) = -2.53$, $p < .05$; t_2 (19) = -3.22, p , .01) there was again no RT difference between the *morphological* and the *consonantal* conditions. In the t-tests for error rates only, the difference between the *morphological* and the *syllabic* condition was significant: error rates in the latter condition were reliably higher than those in the former condition ($t_1(23) = -2.54$, $p < .05$; $t_2(19) = -2.65$, $p < .05$). No other pairwise comparisons were significant. There was a positive correlation of RT with context length in the *consonantal* condition of the monosyllabic item set ($r(20) = .59$, $p < .01$). An ANCOVA with context length as covariate performed separately on the monosyllabic item set revealed a significant effect of context type on RTs ($F_2(2,37) = 3.47$, $p < .05$). Thus, the significant effect of context type obtained in the ANOVA was not exclusively attributable to the lengths of the contexts.

Bisyllabic target words

The result of an ANOVA showed a main effect of context type for RTs ($F_1(2,42) = 7.19$, $p < .01$; $F_2(2,44) = 8.78$, $p < .01$) but not for error rates. Pairwise comparisons of the three conditions showed that the *syllabic* condition was again significantly slower than the other two conditions (*morphological* vs. *syllabic*: $t_1(23) = -3.60$, $p < .01$; $t_2(22) = -3.51$, $p < .01$; *consonantal* vs. *syllabic*: t_1 (23) = -2.92, $p < .01$; $t_2(22) = -3.66$, $p < .01$). As in the monosyllabic item set, there was a positive correlation of context length with RTs in the *consonantal* condition ($r(23) = .56$, $p < .01$). The result of a separate ANCOVA for the bisyllabic item set with context length as covariate was also in line with the monosyllabic analysis: the effect of context type on RTs was still significant when the covariate context length was taken into account ($F_2(2,43) = 6.52$, $p < .01$).

Final embeddings: ***Stop*** items

Although items in this set were a subset of those used in the *fricative* item set, these sets were analysed separately because two different groups of participants took part in each sub-experiment. Raw RTs measured from item onsets were adjusted by subtracting the lengths of the complete items in order to measure RTs from word offsets. Those reactions where subjects pressed the button but did not give adequate vocal responses were treated as errors (2.9%). Two items (*refrein and repliek*) produced error rates of more than 50% and were therefore excluded from further analyses. Table 3.14 and Figure 3.5 show the adjusted RTs and error rates per experimental condition.

The result of an ANOVA showed a highly significant main effect of context type for both RT- and error-analyses (RTs: $F_1(2,66) = 25.87$, $p < .0001$; $F_2(2,54) = 25.66$, $p < .0001$; error rates: $F_1(2,66) = 26.36$, $p < .0001$; $F_2(2,54) = 10.75$,

Table 3.14:
Experiment 3.3: Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions for finally embedded words in the stop item set.

	context		
	consonantal	reduced syllable (Cə)	full syllable (CV)
RT (ms)	669	480	547
Error rate (%)	26	7	12
Example	blepel	selepel	kulepel

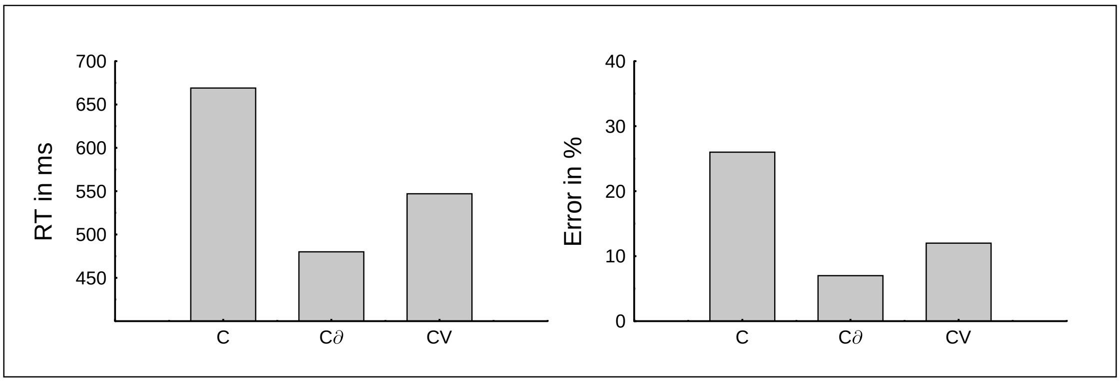


Figure 3.5: Experiment 3.3 - *Final embeddings/Stop items*: Mean RTs (in ms) and mean error rates per experimental condition (C = single consonant context, Cə = reduced syllable context, CV = full syllable context).

$p < .0001$). Pairwise comparisons of the three conditions revealed significant differences between the RTs in all three conditions. Words that were preceded by only a single consonant were spotted slowest as compared to the other two conditions (*consonantal* vs. *reduced syllable*: $t_1(35) = 6.92$, $p < .0001$; t_2 $(27) = 6.53$, $p < .0001$; *consonantal* vs. *full syllable*: $t_1(35) = 3.97$, $p < .0001$; $t_2(27) = 4.52$, $p < .0001$). Interestingly, there was also a significant difference between the two *syllabic* conditions with the RTs in the reduced-syllable condition (Cə) being 67 ms faster than the RTs in the full-syllable condition (CV; $t_1(35) = -3.34$, $p < .01$; $t_2(27) = -2.32$, $p < .05$). T-tests for the error rates showed that the *consonantal* condition elicited reliably more errors than both the *reduced syllable* condition ($t_1(35) = 6.25$, $p < .0001$; $t_2(27) = 4.11$, $p < .0001$) and the *full syllable* condition ($t_1(35) = 4.38$, $p < .0001$; $t_2(27) = 3.02$, $p < .01$).

Although the *reduced syllable* condition produced 5% less errors than the *full syllable* condition, this difference was not significant. Note that these results fully conform with the predictions of the PWC: listeners' reactions are slower and less accurate when words are preceded by a single consonant (i.e., *lepel* in *blepel*)

Table 3.15:
Experiment 3.3: Mean word-spotting latencies (RT, in ms) and Error Rates (%) in each of the three conditions for finally embedded words in the fricative item set.

	context		
	consonantal	reduced syllable (Cə)	full syllable (CV)
RT (ms)	716	532	564
Error rate (%)	28	10	12
Example	blepel	selepel	kulepel

than when they are preceded by a syllable (i.e., *lepel* in either *kelepel* or *kulepel*). Furthermore, the results are in line with what McQueen and Cutler (1998) found: they also reported that the spotting of words embedded in Cə contexts was faster than the spotting of the same words embedded in CV contexts. As the results of a control lexical-decision experiment revealed, this difference was due to acoustic confounds within the target words rather than to the different contexts. This might also be the case for the results of the current experiment.

Final embeddings: *Fricative* items

Again, raw RTs were adjusted by subtracting the lengths of the complete items. Only those responses were taken into account where participants pressed the button and spotted the intended words. A proportion of 2.76% of subjects' responses was excluded because they pressed the button and either failed to give an answer at all or spotted the wrong word. Three items (*repliek, rillen, tonen*) were not included in further analyses because they produced error rates of more than 50%. The adjusted RTs and error rates are listed in Table 3.15 and plotted in Figure 3.6.

As in the *stop* item set, the ANOVA showed a very robust effect of context type on both RTs ($F_1(2,42) = 15.12$, $p < .0001$; $F_2(2,88) = 22.32$, $p < .0001$) and error rates (F_1 $(2,42) = 15.31$, $p < .0001$; $F_2(2,88) = 17.78$, $p < .0001$). The result of a t-test for RTs showed that words embedded in *consonantal* contexts were spotted significantly slower than words preceded by a reduced syllable ($t_1(23) = 3.96$, $p < .0001$; $t_2(44) = 5.87$, $p < .0001$). The same was true for the comparison of the *consonantal* and the *full syllable* condition ($t_1(23) = 3.33$, $p < .01$; $t_2(44) = 4.82$, $p < .0001$). The 32 ms difference between the *reduced syllable* condition and the *full syllable* condition was not statistically reliable. The pattern for error rates was equivalent to the RT pattern: the spotting of words in the *consonantal*

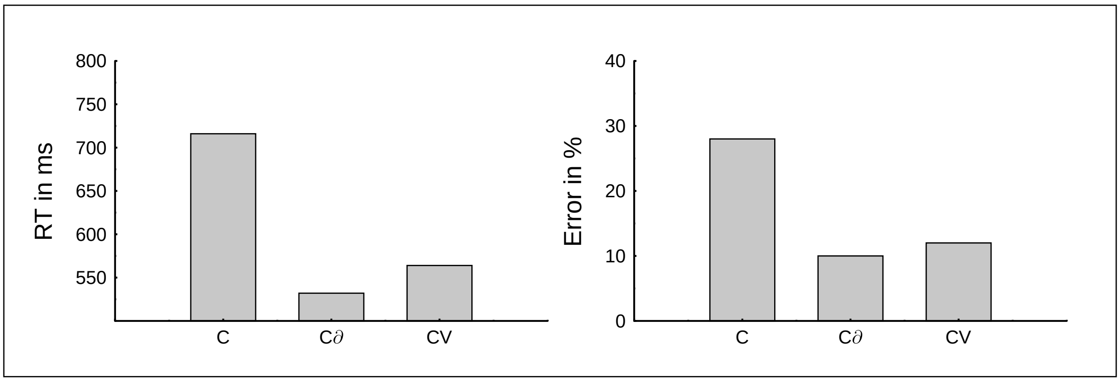


Figure 3.6: Experiment 3.3 - *Final embeddings/Fricative items*: Mean RTs (in ms) and mean error rates (in %) per experimental condition (C = single consonant context, Cə = reduced syllable context, CV = full syllable context).

condition was significantly harder than in either of the other two conditions (*consonantal* vs. *reduced syllable*: $t_1(23) = 4.11$, $p < .0001$; $t_2(44) = 4.96$, $p < .0001$; *consonantal* vs. *full syllable*: $t_1(23) = 3.66$, $p < .01$; $t_2(44) = 4.31$, $p < .0001$). The pairwise comparison of the error rates in the two syllabic conditions was not significant. These results are exactly as the PWC would predict: error rates and RTs are higher in the impossible-word (i.e., the *syllabic*) condition as compared to the two possible-word conditions (i.e., *reduced* and *full syllable* conditions).

Discussion

Experiment 3.3 revealed the familiar pattern for words that were embedded at the onsets of nonsense strings: subjects were slowest in spotting words in the *syllabic* condition where they were expected - according to the PWC - to be fastest. While the pattern of RTs was thus the same as in Experiment 3.1 (and 3.2), the pattern for error rates changed. Those had been in line with the PWC in Experiment 3.1 but reversed in Experiment 3.3. Thus, while errors were lowest in the possible-word (i.e., the *syllabic*) condition in Experiment 3.1, they were now highest in this condition in Experiment 3.3.

However, words that were embedded item-finally (e.g., *lepel* in *selepel*) produced a pattern of results that is perfectly in line with the predictions of the PWC: words embedded in *syllabic* contexts are spotted faster and more accurately than the same words embedded in *consonantal* contexts.

Thus, despite the inclusion of finally embedded targets, listeners still waited until the ends of the nonsense strings when words were embedded item-initially. This strategy therefore proved to be very robust. Making the task harder did not have the desired effect on listeners' performance. They still took their time before they initiated their responses, with the result that the reversed PWC-effect was observed once more. Responses became slower as context length increased. Furthermore, responses to the *syllabic* condition became more errorful than they had been in Experiment 3.1. Thus, not only the RT pattern was the opposite from what was predicted by the PWC but also the error data was reversed. This reversal rules out the earlier hypothesis that longer RTs lead to less errors because of an increased processing time.

However, the proposal that this reversed effect is a task-specific strategy rather than a general segmentation strategy is supported by the results obtained for the finally-embedded target words. Here a robust PWC effect was found in the expected direction. Words in contexts which were possible words of Dutch were detected faster and more accurately than those in impossible-word contexts. Importantly, reduced syllables also seem to count as ’possible words’ in Dutch (replicating earlier findings from McQueen & Cutler, 1998).

Because of the reversed effect in the initially embedded items, there is still no appropriate baseline against which the *morphological* condition might be interpreted. The original question about morphology’s role in segmentation remains unanswered. In the last experiment a final attempt was therefore made to address the role morphemes might play in segmentation. One way to get around the confound of varying lengths in the three relevant conditions is of course to keep that factor constant across the conditions. Because the syllables in the possible word condition could not be shortened, extra syllables were added at the ends of the experimental items in both the *morphological* and the *consonantal* conditions. Thus, sequences like *probleemt* in the *morphological* condition and *probleemp* in the *consonantal* condition were exchanged by *probleemtdaaf* and *probleempdaaf* respectively while the *syllabic* condition was, for example, *probleemdwaaf*.

Note that the segmentation problem stays the same as in the previous experiments because items were created such that the consonants in both the *morphological* and *consonantal* condition belonged to the penultimate syllable of each string. The consonant clusters [td] (as in the *morphological* condition) and [pd] (as in the *consonantal* condition) are not legal onsets of Dutch syllables whereas the cluster [dw] (as in the *syllabic* condition) is. Therefore a syllable boundary was forced by phonotactic constraints to occur after the relevant consonants [t] and [p] in the two former conditions while a syllable boundary in the latter condition was forced after the final phoneme of the embedded word *probleem* because [md] is not a legal coda in Dutch.

Two different outcomes were possible: (A) listeners would still be able to identify the embedded words relatively early (on the basis of their early uniqueness points) but would delay their responses for an equal amount of time (because of the equal lengths of the following syllables) so that no RT differences between the three conditions would be observed, or (B) listeners' performance would be influenced by the factor "(im)possible word" in the three different following contexts - an influence which was masked by the confound of different context lengths in the previous experiments. The first outcome would demonstrate that listeners were indeed able to perform the word-spotting task without the need for segmentation (because of the early uniqueness points) while the second outcome would show that the reversed effects of the previous experiments were only due to the confound of length-differences between the conditions. Thus, in the latter case, a PWC-effect should be observed: RTs in the *syllabic* condition should be faster than those in the *consonantal* condition. Such a result would moreover allow for an interpretation of the morphological case.

Experiment 3.4: Equal context lengths

Method

Materials. The new stimuli were based on the same embedded words as were used in the previous experiments. Only those words were used that had been embedded item-initially in Experiment 3.3. Two major changes in the stimulus construction were carried out: (A) items in both the *morphological* and the *consonantal* conditions were supplemented by additional final syllables that always started with a phoneme that could not form either an onset- or a coda-cluster with the final consonants (e.g., [td] in *probleemtdaaf*, [pd] in *probleempdaaf*) and (B) the onsets of the context syllables in the *syllabic* condition were replaced by CC clusters (i.e., [dw] as in *probleemdwaaf*). The latter constraint was intended to keep the following contexts measured from the final phonemes of the embedded words as constant across the conditions as possible. The contexts were thus equated in length in terms of number of phonemes and syllables. Full lists of the new *stop* and *fricative* experimental items are supplied in Appendix B.

The same target-bearing fillers that had been used in Experiment 3.1 were included in the experiment in order to balance the number of words presented in each context. They were adapted in the same way as the experimental stimuli in the *syllabic* condition. In the *fricative* item set, there were 16 extra target-bearing items (8 mono- and 8 bisyllabic targets), while there were 10 of those in the *stop* item set (3 mono- and 7 bisyllabic targets).

The filler items (that had no embedded words) were based on those of previous experiments but were adapted so that they matched the experimental stimuli structurally. Thus, fillers that had matched experimental items in the *morpholog-*

ical or *consonantal* conditions of Experiment 3.3 were supplemented by syllables that were chosen such that they had onsets that could not form coda- or onset-clusters with the final consonants of the filler items. The onsets of the final syllables of those fillers that had matched experimental stimuli in the *syllabic* condition of Experiment 3.3 were exchanged by CC clusters.

As in the earlier experiments, three different lists were created for each experiment with the contexts for each target word counterbalanced across lists. In the *fricative* experiment each list contained 16 experimental items in each condition while in the *stop* experiment there were only 10 experimental items per condition. Each subject was presented with only one of these lists so that none of them heard any target more than once. Each list in each item set contained all fillers (128 in the *fricative* item set and 80 in the *stop* item set).

The lists therefore differed only in terms of the contexts of the target-bearing items. Two random orders of each list were then made, in which only the order of stimulus presentation differed between these lists. The randomizations were constrained in the following way: between two target-bearing strings there was at least one filler but there were never more than 3 fillers between experimental items. A total of six differently randomized item lists was thus produced for each experiment. Within each list the target positions were kept constant.

Recording. The items were recorded by a female Dutch native speaker of Dutch (who was not the same as in the previous experiments). The recording was made in an sound-attenuated booth with the same recording facilities as in the previous experiments. The stimuli were redigitized from DAT onto a computer and were then measured and spliced into individual speech files, using the Xwaves/ESPS speech editor.

Subjects and Procedure. For each experiment, thirty-six students from Nijmegen University were recruited from the Max Planck Institute’s subject pool and were paid for their participation. None of them reported any hearing deficit. The instructions and procedure were exactly as in the previous word-spotting experiments. After a brief practice session the main experiment was divided into two parts allowing subjects one pause in between. Sessions in the *fricative* experiment were slightly longer than those in the *stop* experiment but lasted no longer than 25 minutes.

Table 3.16:
Experiment 3.4: Mean word-spotting latencies (RT, ms), Error Rates (%), and mean context lengths (in ms) in each of the three conditions for the bisyllabic targets in the stop item set.

	context		
	morphological	consonantal	syllabic
RT (ms)	445	458	357
Error rate (%)	11	8	1
Mean context length (ms)	569	575	500
Example	probleemtdaaf	probleempdaaf	probleemdwaaf

Results and Discussion

Stop items

As in all previous experiments, the raw RTs were adjusted by subtracting the lengths of the embedded words in order to yield RTs from word offsets. A proportion of 4% of all responses was excluded from further analyses because subjects either failed to give a response at all or spotted the wrong word but nevertheless had pressed the button. There were six items (*teneur, nul, tor, wang, geul, bil*) that produced error rates of more than 50% and were thus - as in previous experiments - excluded from further analyses. There was one item (*riool*) which was never spotted by any subject in the *consonantal* condition and which was therefore also excluded from further analyses. A closer inspection of the subjects' error rates revealed that 23 out of 36 subjects (64%) failed to spot monosyllabic target words in at least one of the three conditions. The four remaining monosyllabic target words were therefore excluded from further analyses as well. Only the bisyllabic item set was analysed. See Table 3.16 and Figure 3.7 for mean RTs and error rates for bisyllabic target words in the three conditions.

The result of an ANOVA on the bisyllabic item set showed a reliable context effect for RTs ($F_1(2,66) = 12.78$, $p < .0001$; $F_2(2,36) = 6.23$, $p < .01$) and a weaker effect in the error rates ($F_1(2,66) = 13.02$, $p < .0001$; $F_2(2,36) = 2.69$, $p < .09$), where only the subjects' analysis was significant. In individual t-tests the *syllabic* condition produced reliably faster RTs than both the *morphological* ($t_1(35) = 3.62$, $p < .001$; $t_2(18) = 2.67$, $p < .02$) and the *consonantal* ($t_1(35) = 4.93$, $p < .0001$; $t_2(18) = 3.33$, $p < .01$) conditions. There was no significant difference between the latter two conditions. A pairwise comparison of the error rates in the three conditions again showed significant effects in the subjects' analyses but not in the items' analyses. Responses in the *syllabic* condition were more

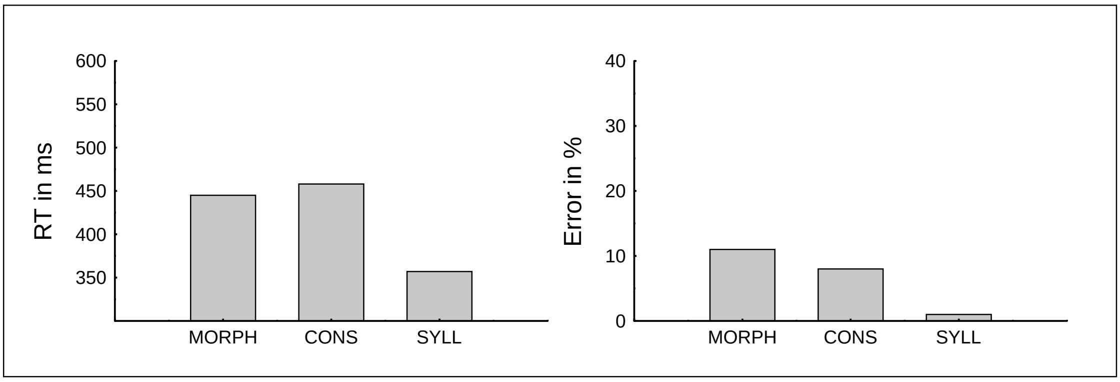


Figure 3.7: Experiment 3.4 - *Stop items*: Mean RTs (in ms) and mean error rates (in %) per experimental condition (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

accurate than responses in the other two conditions (*syllabic* vs. *morphological*: $t_1(35) = 4.96$, $p < .0001$; *syllabic* vs. *consonantal*: $t_1(35) = 3.57$, $p < .001$), while the *morphological* and the *consonantal* conditions did not differ from each other.

Fricative items

In Table 3.17 the adjusted mean RTs and error rates are listed per experimental condition for the *fricative* item set (see also Figure 3.8). Those trials (3% of all trials) where subjects pressed the button but failed to spot the correct word or did not make a vocal response at all were excluded from further RT analyses. Seven items (*parool, ham, wang, bil, rol, gong, tang*) were excluded from any analyses because of error rates higher than 50%. None of the subjects spotted the item *vorm* in the *morphological* condition so that this item was also excluded from the analyses. Furthermore, six subjects were not included in any of the analyses because they failed to spot target words in at least one of the experimental contexts.

An overall ANOVA revealed a significant effect of context type for both RTs ($F_1(2,54) = 9.36$, $p < .0001$; $F_2(2,74) = 8.03$, $p < .001$) and error rates ($F_1(2,54) = 53.22$, $p < .0001$; $F_2(2,74) = 25.31$, $p < .0001$). Also the factor syllable number was significant in both analyses (RTs: $F_1(2,27) = 170.78$, $p < .0001$; $F_2(2,37) = 74.93$, $p < .0001$; error rates: RTs: $F_1(2,27) = 55.11$, $p < .0001$; $F_2(2,37) = 37.31$, $p < .0001$). Neither in the RT nor in the error analyses was there an interaction of these factors. A t-test for RTs showed a very similar pattern as in the *stop* item set. Subjects were significantly faster in spotting targets embedded in *syllabic* contexts than they were in spotting targets embedded in either *consonantal* ($t_1(29) = 2.82$, $p < .01$; $t_2(38) = 3.95$, $p < .0001$) or *morphological* contexts although the item analysis was only marginally significant in the latter

Table 3.17:
Experiment 3.4: Mean word-spotting latencies (RT, ms), Error Rates (%), and mean context lengths (in ms) in each of the three conditions for the fricative item set.

	context		
	morphological	consonantal	syllabic
RT (ms)	486	492	436
Error rate (%)	9	26	7
Mean context length (ms)	524	518	470
Examples	forelsdur/ deursbug	forelfdur/ deurkbug	foreldrur/ deurblug

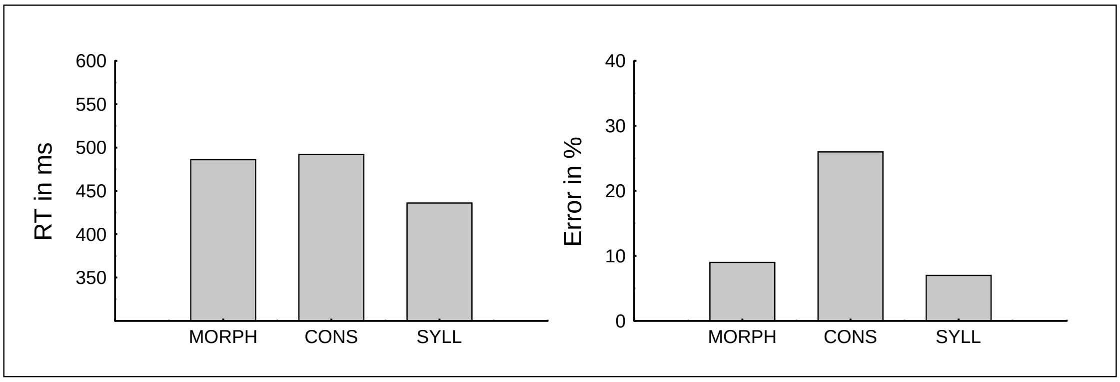


Figure 3.8: Experiment 3.4 - *Fricative items*: Mean RTs (in ms) and mean error rates (in %) per experimental condition (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

comparison ($t_1(29) = 2.78$, $p < .01$; $t_2(38) = 1.89$, $p < .07$). A somewhat different pattern was found for the error rates: subjects made significantly more errors when they had to spot words in *consonantal* contexts as compared to the other two contexts (*consonantal* vs. *morphological*: $t_1(29) = -6.08$, $p < .0001$; $t_2(38) = -3.57$, $p < .01$; *consonantal* vs. *syllabic*: $t_1(29) = 7.4$, $p < .0001$; $t_2(38) = 4.66$, $p < .0001$). The item set was analysed separately for mono- and bisyllabic target words. The respective RTs and error rates are listed in Table 3.18.

An ANOVA calculated individually for the *monosyllabic* items showed a reliable effect of the factor context type on both RTs and error rates (RTs: ($F_1(2,54) = 5.28$, $p < .01$; $F_2(2,30) = 4.7$, $p < .02$; error rates: ($F_1(2,54) = 51.66$, $p < .0001$; $F_2(2,30) = 18.79$, $p < .0001$). In a t-test, RTs in the *syllabic* condition were reliably faster than in the *consonantal* condition ($t_1(29) = 3.32$, $p < .01$; $t_2(15) = 3.32$, $p < .01$). The RT difference between the *syllabic* and the *morphological* condition was only significant by subjects but not by items ($t_1(29) = 2.11$, $p <$

Table 3.18:
Experiment 3.4: Mean word-spotting latencies (RT, ms), Error Rates (%), and mean context lengths (in ms) in each of the three conditions for the fricative item set listed separately for mono- and bisyllabic target words.

monosyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	644	701	568
Error rate (%)	16	51	11
Mean context length (ms)	543	558	485
Example	deursbug	deurkbug	deurblug

bisyllabic targets			
		context	
	morphological	consonantal	syllabic
RT (ms)	389	412	353
Error rate (%)	5	8	4
Mean context length (ms)	513	503	461
Example	forelsdur	forelfdur	foreldrur

.05; $t_2(15) = 1.53$, n.s.). Although there was a 56 ms difference between the *morphological* and the *consonantal* condition this difference was not reliable in a pairwise comparison. T-tests calculated on the error rates showed that the *consonantal* condition produced reliably higher error rates than the other two conditions (*consonantal* vs. *morphological*: $t_1(29) = -7.07$, $p < .0001$; $t_2(15) = -3.88$, $p < .001$; *consonantal* vs. *syllabic*: $t_1(29) = 8.86$, $p < .0001$; $t_2(15) = 5.8$, $p < .0001$).

The effect of the factor context type on both RTs and error rates in the bisyllabic item set was only significant in the subject analysis (RTs: ($F_1(2,54) = 6.7$, $p < .01$; $F_2(2,44) = 2.57$, n.s.; error rates: ($F_1(2,54) = 3.08$, $p < .05$; $F_2(2,44) = 2$, n.s.). Pairwise comparisons revealed that RTs in the *consonantal* condition were reliably longer than in the *syllabic* condition ($t_1(29) = 3.08$, $p < .01$; $t_2(22) = 2.34$, $p < .05$). Similarly, RTs in the *morphological* condition were longer than in the *syllabic* condition although this was only true for the subject analysis ($t_1(29) = 2.1$, $p < .05$; $t_2(22) = 1.12$, n.s.). T-tests that compared the error rates of the three conditions did not reveal any significant results. There was only a vague tendency towards lower error rates in the *syllabic* condition as compared to the *consonantal* condition ($t_1(29) = 1.6$, n.s.; $t_2(22) = 2.0$, $p < .06$).

At first sight, the additional manipulation of the different context lengths seemed to have the desired effect on word-spotting performance. When context lengths

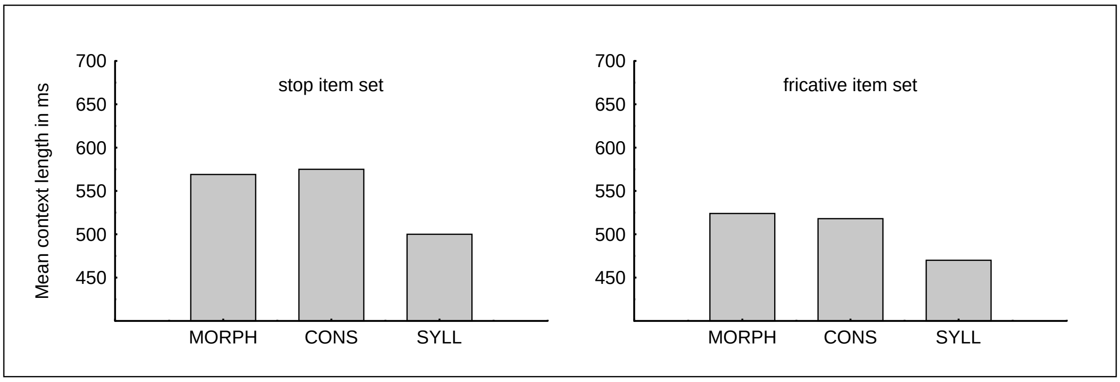


Figure 3.9: Experiment 3.4: Mean context lengths (in ms) per experimental condition in each item set (MORPH = morphological context, CONS = consonantal context, SYLL = syllabic context).

were kept constant across the three different conditions, the predicted PWC effect was observed: the spotting of words in a possible (i.e., *syllabic*) context was faster than in an impossible (i.e., *consonantal*) context. But before the results can be further interpreted with respect to the *morphological* condition it needs to be established that the context lengths in the three different conditions indeed did not differ. The mean lengths of following contexts for the bisyllabic *stop* items are listed in Table 3.16 and plotted in Figure 3.9.

An ANOVA performed on the context lengths revealed a significant effect of the factor context type ($F_2(2,36) = 14.38$, $p < .0001$). As was shown by t-tests, this effect was due to shorter contexts in the *syllabic* condition as compared to contexts in both the *morphological* and the *consonantal* conditions (*syllabic* vs. *morphological*: $t_2(18) = 4.0$, $p < .001$; *syllabic* vs. *consonantal*: $t_2(18) = 4.46$, $p < .0001$). No such difference was observed between the *morphological* and the *consonantal* conditions. Furthermore, the result of a correlational analysis showed a highly significant positive correlation of RT with context length in the *morphological* condition ($r(19) = .35$, $p < .01$).

In Table 3.17 the respective context lengths are listed for the *fricative* item set (see also Figure 3.9). Exactly the same result was obtained as in the *stop* item set: there was a main effect of context type on the lengths of the different contexts ($F_2(2,74) = 50.32$, $p < .0001$) while the factor syllable number and the interaction of both factors was not significant. T-tests again showed that the context length in the *syllabic* condition was significantly shorter than the lengths in the other two conditions (*syllabic* vs. *morphological*: $t_2(38) = 7.36$, $p < .0001$; *syllabic* vs. *consonantal*: $t_2(38) = 7.73$, $p < .0001$). There was again no difference between the context lengths in the *consonantal* as compared to the *morphological* condition. Furthermore, there was a highly significant correlation of RT with context length in the *consonantal* condition ($r(39) = .43$, $p < .01$).

When the mean context lengths of mono- and bisyllabic target words are considered separately, they reveal the same patterns. Table 3.18 lists the mean context lengths for mono- and bisyllabic target words per condition. Separate ANOVAs revealed significant context effects in both item sets (monosyllabic items: $F_2(2,30) = 22.54$, $p < .0001$; bisyllabic items: $F_2(2,44) = 28.13$, $p < .0001$). Pairwise comparisons of the context lengths in the three conditions showed that in both item sets the mean lengths in the *syllabic* condition were significantly shorter than in the other two conditions. There were no differences between the mean context lengths in the *morphological* and *consonantal* conditions for either mono- or bisyllabic items (monosyllabic items / *syllabic* vs. *morphological*: $t_2(15) = 3.84$, $p < .01$; *syllabic* vs. *consonantal*: $t_2(15) = 6.97$, $p < .0001$; bisyllabic items / *syllabic* vs. *morphological*: $t_2(22) = 7.22$, $p < .0001$; *syllabic* vs. *consonantal*: $t_2(22) = 4.75$, $p < .0001$). A positive correlation of RTs with context lengths was only marginally significant in the *consonantal* condition of the monosyllabic item set ($r(16) = .49$, $p < .06$), while there were no correlations in either of the conditions in the bisyllabic item set.

This pattern of results suggests that the robust context effects found in RTs might be mainly due to differences in the context lengths across the three conditions rather than to a segmentation process based on the PWC. The correlational analyses revealed positive correlations of RTs with context lengths in all except one subset of items. Thus, the shorter the context following an embedded word, the faster subjects initiated their responses. In order to test whether the context effects on RTs were exclusively due to the factor context length, the latter factor was used as a covariate in ANCOVAs calculated for all subsets of items (ANCOVAs in the *fricative* item set had to be calculated separately for mono- and bisyllabic item sets because the number of items differed between the item sets). All three ANCOVAs revealed the same results: as soon as the factor context length was included in the analyses, the context effect on RTs vanished (bisyllabic *stop* items: $F_2(2,35) = 1.95$, n.s.; monosyllabic fricative items: $F_2(2,29) = 0.93$, n.s.; bisyllabic *fricative* items: $F_2(2,43) = 1.97$, n.s.). This result convincingly shows that subjects in this study used the same strategy as in previous experiments: they waited until the ends of nonsense sequences before they initiated their responses. Because the contexts in Experiments 3.1 - 3.3 were always shorter in the *consonantal* and the *morphological* conditions, shorter RTs were obtained for these conditions. Although an attempt was made to control for this factor in the latest experiment, this control did not succeed. Context lengths again varied between the conditions, but this time in the opposite direction than in the other experiments. As a consequence of shorter context lengths in the *syllabic* condition, the RTs were also shorter in this condition.

General Discussion

The experiments reported in Chapters 2 and 3 were designed to examine the role morphology might play during the segmentation of spoken language. The error data of Experiment 2.1A indicated that morphemes might not play a different role in segmentation than morphologically meaningless consonants. The results of Experiment 2.1A were, however, not entirely unambiguous since listeners' RTs in the *morphological* condition were between the fastest (i.e, *syllabic*) and the slowest (i.e., *consonantal*) conditions. These differences were however not significant. There was furthermore an indication of a potential difference between the two different item sets (*stop* and *fricative* items). In the second word-spotting experiment (Experiment 3.1) the two item sets were therefore presented in two separate item lists. This separation, moreover, made it possible to include more stimuli, with the advantage of an increase in experimental power. The hope was that the pattern observed for the morphological case in Experiment 2.1A would become clear in Experiment 3.1. However, what was observed instead was a reversal (in RTs) of the PWC effect as established by Norris et al. (1997). The condition where segmentation and therefore word-spotting was predicted to be easiest (i.e., the *syllabic* condition) produced the slowest RTs. The spotting of *kopie* 'copy' in *kopiekel* was thus significantly slower than the spotting of the same word in either *kopiet* (i.e., *morphological* context) or *kopiek* (i.e., *consonantal* context). According to the PWC, the opposite result should have been found: the spotting of words in a possible-word (i.e., *syllabic*) context should be easier than the spotting of a word in an impossible-word (i.e., *consonantal*) context because the parser should segment speech such that no impossible words are left over.

The error rates in this condition, on the other hand, agreed with the principles of the PWC: they were lower than in the impossible word condition. The *morphological* condition was in some sense still between the other two conditions: while RTs in this condition were the same as in the impossible word condition the opposite was true for error rates, which were similar to those in the possible word condition. The results were similar across the *stop* and *fricative* item sets. Due to the missing baseline (i.e., the comparison of the *syllabic* and the *consonantal* conditions), it was still not possible to draw a strong conclusion about the status of morphemes in segmentation. Because in both item sets the RT reversal could have been driven by an acoustic confound in the prosodic structure of the bisyllabic item sets, a control experiment (Experiment 3.2) was conducted which tried to eliminate that confound. The same pattern of results was obtained as in Experiment 3.1. Correlations of RTs with the lengths of following contexts suggested that listeners in these experiments employed a waiting strategy instead of responding as fast as possible. Due to the early uniqueness points of the targets, subjects had been able to identify the embedded words before they had heard the following contexts. As a result of that early recognition, no further segmentation was required to solve the word-spotting task. But although listeners were able to identify the words early and therefore could have responded early, they nevertheless decided to wait until the ends of nonsense strings before they initiated their responses.

Remember that the correlations of context length with RTs and error rates in Experiment 2.1 (with only monosyllabic targets) were not significant and that the results in this experiment were in the predicted directions. The waiting strategy observed in Experiments 3.1 - 3.4 seems therefore to be mainly due to the bisyllabic target words. This strategy then “spilled over” from the bisyllabic to the monosyllabic words.

Experiments 3.3 and 3.4 attempted to deal with the problems which led to the waiting strategy but showed in principle the same results as the previous experiments. While in Experiment 3.3 the condition that was predicted to be easiest (i.e., the *syllabic* context) again produced the slowest responses, the opposite result was obtained in Experiment 3.4. But correlational analyses and analyses of covariance showed that the results of Experiment 3.4 could again be attributed to the factor context length rather than to the factor context type. Note, however, that the results obtained for the finally-embedded target words in Experiment 3.3 (e.g., ’lepel’, *spoon* in ’blepel’ vs. ’kulepel’) were in line with the predictions of the PWC: possible-word contexts produced reliably lower RTs and error rates than impossible-word contexts. The reversed PWC effect was therefore confined to words that were embedded item-initially. Furthermore, the results for finally-embedded targets convincingly demonstrated that the experiments had sufficient power to pick up PWC effects.

Unfortunately, the pattern of error rates across the experiments was far from consistent. However, only Experiment 3.3 revealed a reversed PWC effect for errors in item-initially embedded targets. In all other experiments, there were either no significant differences between the error rates in the three conditions or there were significant differences in the predicted direction (i.e., less errors in the *syllabic* condition than in the *consonantal* condition). Thus, the reversal of the predicted PWC effect seems to be mostly confined to RTs, which is consistent with the idea of a waiting strategy.

What is consistent across all experiments, furthermore, is the difference that was always observed between the mean RTs and mean error rates for monosyllabic items as compared to bisyllabic items irrespective of contexts. In almost all ANOVAs (on both RTs and error rates) there was a strong effect of the factor number of syllables which was due to the fact that bisyllabic target words were always reliably longer than monosyllabic targets. This had an effect on both dependent variables. The longer the words were, the more processing time subjects had and the faster and more accurate they were in their performance. Overall it was harder for subjects to spot monosyllabic words as compared to bisyllabic words.

This further supports the hypothesis that the early uniqueness points of the bisyllabic target words enabled the subjects to perform the word-spotting task without the need for segmentation. This is also consistent with the results of the first experiment, in which only monosyllabic target words were used and where the PWC effect was observed. What remains to be explained, though, is why monosyllabic target words in the subsequent experiments showed very similar RT patterns to those of the bisyllabic target words. One explanation might be that the waiting strategy was initiated by the bisyllabic items but was then also applied to monosyllabic items. It could be that subjects indeed needed to segment the nonsense strings in order to solve the task for the monosyllabic targets, but that an effect of graded segmentation difficulty was masked by the length effect. Because there were no bisyllabic targets in the very first experiment, listeners did not employ a waiting strategy and thus the classical PWC effect was obtained. Only when bisyllabic words were included in the experimental lists did the effect reverse.

Other results from English (Norris et al., 1997; Norris et al., 2001) and Dutch (McQueen & Cutler, in preparation), however, suggest that the inclusion of bisyllabic target words may not be necessary to produce a data pattern like that observed in Experiments 3.1 - 3.4. Norris et al. (1997, 2001) show that the inclusion of bisyllabic target words in the experimental lists does not necessarily result in a waiting strategy. In each of these sets of experiments, both mono- and bisyllabic target words were used. The results were in line with the predictions of the PWC: RTs were faster in *syllabic* than in *consonantal* contexts. This was true for both item-initially and item-finally embedded words. Moreover, the authors do not report any major differences between mono- and bisyllabic targets. Furthermore, McQueen and Cutler (in preparation) conducted a word-spotting experiment in Dutch where only monosyllabic target words were used. Targets were presented either item-initially or item-finally. The results for the monosyllabic target words that were embedded at the onsets of nonsense strings showed the same RT pattern as the current experiments: listeners' responses were slowest in the condition that was predicted to be fastest (i.e., targets in following syllabic contexts). Targets embedded at the ends of nonsense strings, however, produced a reliable PWC effect (i.e., faster responses to targets in *syllabic* than in *consonantal* contexts). For item-initially embedded targets, as in Experiments 3.1 - 3.4, there was also a positive correlation of RTs with context length, suggesting that the subjects in the McQueen and Cutler study also waited until the ends of the whole items before they initiated their responses. If there were no bisyllabic targets that initiated a waiting strategy in this experiment, what other factor might have been responsible for the reversed PWC effect?

It was argued in Chapter 2 that the confound of *sequential probabilities* might cause the PWC effect to vanish. Those monosyllabic targets that had been used in the *fricative* item set of Experiment 2.1A did not produce a PWC effect: subjects were as fast in spotting words in *consonantal* contexts (e.g., *duim* 'thumb' in *duimf*) as they were in spotting words in *syllabic* contexts (e.g., *duim* in *duimfoel*). It was argued that this lack of a significant difference was mainly due to the low sequential probabilities leading up to the final phonemes in the *consonantal* condition in the *fricative* item set of Experiment 2.1A (e.g., [mf] in *duimf*). If a coda-cluster is very unlikely in a given language, listeners might be able to use that information to place word boundaries at the offset of the target words. This information might be strong enough to overrule the principles of the PWC. Similarly, the sequential probabilities in the *consonantal* condition of the McQueen and Cutler (in preparation) study were quite infrequent and may therefore have caused the disappearance of the predicted effect. In contrast to the Dutch experiment, the sequential probabilities in the original English study (Norris et al., 1997) were much higher.

But note that the **disappearance** of the effect - either due to early uniqueness or due to low sequential probabilites - is not the same as a **reversal**. It is quite clear from the current results that the reversal of the predicted effect reflects some sort of waiting strategy. Why listeners in Experiment 3.1 - 3.4 employed such a waiting strategy remains uncertain. But it appears to be the case that as soon as the task can be solved on the basis of information other than the principles of the PWC, listeners can start to employ this waiting strategy: even if the word could have been recognized early (on the basis of either early uniqueness or low sequential probabilities), listeners nevertheless waited until they heard the whole string before they initiated their responses, as if they wanted to double-check word-recognition. Note, however, that whether or not listeners choose to adopt this strategy appears to depend on a fine balance of different features. For example, although the sequential probabilities in a subset of the items used in Experiment 2.1A were low, and targets were therefore relatively easy to spot irrespective of the type of following context, the task was apparently still difficult enough - due to the short monosyllabic words and the high probabilities in the *stop* item set - to reveal an effect that was in line with the predictions of the

PWC. Since the sequential probabilities were infrequent across all items in the McQueen and Cutler (in preparation) study (which also used only monosyllabic targets), the task became easy enough for listeners (a) to segment the input without processing the context following the target and (b) to employ a waiting strategy.

Although the sequential probabilities were higher for all items in Experiment 3.1, the task overall became so easy - due to the early uniqueness of the bisyllabic targets - that listeners in some sense had the luxury to use a waiting strategy. Furthermore, the early uniqueness of the bisyllabic targets - that were used in all experiments except Experiment 2.1 - appeared to be so helpful that even higher task demands could not delay the identification of those words.

It might, however, be possible to use another experimental design to prevent subjects from using a waiting strategy. Imagine that the following contexts that were used in Experiment 3.4 (e.g., *deur* in *deursbug*) were to be extended by yet another syllable (e.g., *deur* in *deursbugfum*). This might make a waiting strategy (which probably was unconscious in the current study) so obvious and unreasonable that listeners might choose to respond as fast as possible rather than to wait until they have heard the last phoneme of the string. Motivating listeners to respond as fast as possible might give the following contexts a chance to constrain or facilitate the word-spotting task in a way that is predicted by the PWC.

Apart from the various confounds reported for the different experiments, one other finding proved to be very consistent throughout the experiments. The RTs in the *morphological* condition always clustered with those in the *consonantal* condition (apart from in Experiment 2.1A). One therefore might be tempted to conclude that inflectional morphemes do not have a special status for the segmentation process. This would furthermore imply that there have to be full form access representations for morphologically complex verbs in order to allow for a smooth segmentation process. Forms like *loop, loopt, loper, lopend* etc. would thus all be activated and compete for recognition. Because the inflected form *loopt* would have its own access representation, the activation of the word would not be penalized by the PWC. If there were no such full form representations of inflected words, the PWC would hinder the recognition process each time it encountered an inflected verb like *loopt*. The word *loop* would be activated, but its activation would soon be reduced because the PWC would encounter the single consonant *-t*, which is not a possible word in Dutch. This would delay if not prevent recognition of *loop-t*. In other words, if there was no access representation for *loopt* and the PWC did not treat a morphological '-t' differently from other single consonants, the parser would have difficulty recognizing *loopt*.

This interpretation, however, is not sufficiently supported by the data. The results suggest that segmentation difficulty was not measured sufficiently in the current experiments. This conclusion is based on the observation that RTs in Experiments 3.1 - 3.4 depended almost exclusively on the *lengths* of the following contexts and not on the *type* of those contexts. It is likely that context-*type* was not effective in these experiments because of one feature of the bisyllabic target words: early uniqueness. If, instead of segmentation difficulty, a task-specific strategy was measured, one cannot draw any conclusions about a theoretical question that focused on *segmentation*. The question about the role of morphology during segmentation in Dutch must therefore remain unanswered for the time being.

I think it might be possible, however, to test the same question in another language, where the construction of stimuli might be less constrained than in Dutch. The words used in these experiments were the only ones that met all the constraints, but carried the confound of early uniqueness. If it were possible to find enough words with late uniqueness points in German, for example, very similar experiments could be conducted in German. Future research might therefore be able to answer the question about the relationship between morphology and segmentation.

The role of morphology in phonetic decision-making - Part I

Chapter 4

Introduction

Very early in the process of decoding spoken language the listener is confronted with the task of identifying the segments that make up the continuous speech stream. The identification of phonemes is not only driven by the acoustic signal itself. There is now an impressive amount of evidence that the identification of speech sounds is strongly influenced by the linguistic context these segments are produced in (Warren, 1970; Sawusch & Jusczyk, 1981; Samuel, 1997). In certain circumstances linguistic context alone can determine the identification of sounds if appropriate acoustic information is absent. In 1970 Warren demonstrated that listeners automatically and unconsciously restored segments that had been replaced by non-linguistic sounds like coughs (the phoneme restoration effect). This restoration must have been driven exclusively by the context since the acoustic information was non-linguistic in nature. The relative contribution of acoustic information on the one hand and contextual information on the other hand to phoneme identification has been at the center of psycholinguistic research for more than two decades now.

In order to resolve the debate about exactly how and when these two factors modulate phoneme identification, researchers have extensively looked at acoustic-phonetic ambiguities. More accurately, they have investigated to what extent listeners use different sources of information to resolve these ambiguities when they are explicitly asked to identify ambiguous sounds as one or the other phoneme. The phonetic categorization task has been frequently used in those kinds of studies. In phonetic categorization experiments, subjects are typically asked to label sounds that vary on an acoustic continuum between two phonetic categories in respect to one phonetic feature. The sounds that the subjects are asked to label are presented in some sort of context that is supposed to influence the subjects' performance. Classically, higher-order influences on such a categorization task appear as shifts in the categorization function towards one or the other endpoint depending on the lexical status of the target string or the syntactic or semantic appropriateness of the target in a sentence context. This shift is usually confined to the boundary region of the continuum (i.e., the most ambiguous stimuli of the continuum).

The manipulated sounds in these studies - which I will discuss in more detail below - have always been phonemes that were part of uninflected content words or function words (Borsky, Tuller, & Shapiro, 1998; Connine, 1987; Ganong, 1980; Miller, Green, & Schermer, 1984). It has never been the case that a phoneme was manipulated that itself was also a meaningful unit in morphological terms. An interesting question is thus whether and how far the linguistic (sentential) environment can bias the perception of inflectional morphemes. Are listeners more likely to label a sound ambiguous on a place-of-articulation continuum as [t] (3rd Ps. Sg. marker for Dutch verbs) if the sound was presented at the end of a syntactically predictable *verb* than if the sound was part of a (contextually) predictable *noun*?

Imagine the two different small phrases in Dutch “de tante gaat” (*the aunt walks*) vs. “een brede straat” (*a wide street*). In the first example the phoneme [t] is an inflectional morpheme and is required in word final position in order to make the sentence grammatical. The [t] is therefore not only predictable on the basis of the context because it forms a lexical item, but also is required by the syntax. This is not true for the second example where the [t] is only needed in order to yield a lexical element. This sentence would not be ungrammatical if the last phoneme was not a [t] in the sense the first one would be. Would such a syntactic predictability reflect itself in a different categorization function for verb phrases (VP) than for noun phrases (NP) if listeners were required to label ambiguous sentence-final sounds? If the final sound varied from [t] to [k] (VP: *de tante gaat - de tante gaak*; NP: *een brede straat - een brede straak*), would there be differences in the way listeners label the final sounds? Note that in both cases the [k] endpoints (*gaak* and *straak*) form nonwords. Such an outcome could have interesting implications for theories of morphological representation and processing. I will discuss these more explicitly later in this Chapter.

Sentence context effects

The current study was designed to investigate potential sentential influences on the processing of inflectional morphemes. It was based on earlier research on sentential influences on phonetic categorization. Miller et al. (1984), for example, showed that listeners’ identification of sentence-final ambiguous words was strongly biased by the semantic contexts these ambiguous words were presented in. When they heard a sentence frame like “She needs hot water for the ...” with the final element being ambiguous between *bath* and *path*, they tended to label the word as *bath*. The same ambiguous element was identified as *path* when the preceding sentence was “She liked to jog along the ...”. However, these sentential context effects were not mandatory. The semantic context only influenced the selection of the final word when listeners’ attention was explicitely drawn to sentential information by asking them to judge not only the final word but also the sentence frame.

Similarly, Connine (1987) reported that subjects labeled sounds ambiguous on a voice-onset-time (VOT) continuum between [t] and [d] so that they formed the words (either *dent* or *tent*) which were appropriate for the sentential context. Furthermore, she found that RTs were faster for sentence-consistent answers than for inconsistent answers. This time benefit was however observed for reactions to endpoint stimuli but not to stimuli in the boundary region. On the basis of this RT pattern, Connine (1987) concluded that sentential semantics could influence the decision about an ambiguous sound but could not, as it were, preactivate a semantically-consistent lexical entry. She therefore attributed the time benefit for consistent reactions at continuum endpoints to an integration difficulty for inconsistent stimuli. This integration was more time consuming as compared to the integration of consistent lexical items. Because in the ambiguous region the identification of the phoneme is itself time consuming, the consistency effect cannot exert its influence on the speed of the decision process.

Contradictory results were obtained by Borsky et al. (1998), who also presented semantically-biasing sentence contexts with ambiguous words embedded within them. The first phoneme of the target words varied on a voice-onset-time (VOT) continuum between [g] and [k] and listeners were asked to judge whether a visually-presented probe at the offsets of target words corresponded to what they had just heard (e.g., *The laughing dairyman hurried to milk the [?ot] in the drafty barn*). Target words were either *goat* or *coat*. Borsky et al. (1998) reported a boundary shift in favor of the biasing context in the ambiguous region. In contrast to the Connine (1987) results, however, there was an RT benefit for identification of stimuli in the ambiguous region when the decisions were congruent with the preceding context. No such RT benefit was observed for reactions to continuum endpoint stimuli. The authors therefore claimed that sentential context influences very early processes of phonological encoding rather than processes of decision making. Aside from the controversy about the locus of sentence context effects (which will be discussed in more detail later), these results nicely demonstrate that preceding semantic context influences listeners’ identification of ambiguous sounds. It is important to note that this influence only shows its effect in situations where listeners cannot rely on the acoustic signal alone to solve the task. As soon as the acoustic-phonetic information is unequivocal (as is the case at the endpoints of continua) listeners' phoneme identification is not influenced by sentential contexts any more. This can be concluded because phonemes at continua endpoints are uniformly labeled correctly even if these labels sometimes lead to contextually inconsistent decisions.

Interestingly, not only can preceding sentential context influence the decisions listeners make to ambiguous sounds. Connine, Blasko, and Hall (1991) demonstrated that listeners' identification performance is also influenced when biasing sentential context follows the ambiguous word within a certain time window. In this study, Connine and colleagues asked subjects to identify the first sound of the third word in each sentence. These sounds varied in VOT on a five step continuum of which the two endpoints formed the words *dent* and *tent*. Sentence onsets were kept neutral and biasing context was made available only after the critical words. Furthermore, Connine et al. (1991) varied the temporal interval within which biasing information was presented. Three different conditions were contrasted: disambiguating information was presented immediately after the critical word, or the distance between the critical word and disambiguating context was increased either to three syllables or from three to six syllables. The crucial question was how long listeners would delay their decisions in order to benefit from contextual information. The results suggest that subsequent context can influence phoneme identification only in a limited time window. Connine et al. (1991) found a typical shift of the categorization function in accordance with biasing context when disambiguating information was presented in a time window of three syllables after the critical word. This shift was only observed in the ambiguous region but not at continuum endpoints. If, however, the relevant information was made available only on the sixth syllable after critical word offset, listeners decisions were not influenced by that information. When they compared the immediate condition with the three syllable delay condition, the authors found no difference between the two categorization functions. This means that the context effect was not stronger when disambiguating information was available immediately than when it was delayed for three syllables.

In contrast to the rather well-documented effect of sentence semantics on phonemic decision-making (see also Samuel, 1981), syntactic influences on phonetic categorization have been investigated comparatively rarely (Isenberg, Walker, & Ryder, 1980; van Alphen & McQueen, 2001). In a recent study, van Alphen and McQueen (2001) looked at syntactic influences on the identification of Dutch function words. Van Alphen and McQueen created a voicing continuum of which the voiced endpoint formed the Dutch article [də] while the voiceless endpoint was the infinite marker [tə]. As in the Connine et al. (1991) study, disambiguating information was provided after the critical word. They created three different context sentences with one requiring the function word [də] (i.e., "We proberen de schoenen" *We try the shoes*), the other one requiring [tə] (i.e., "We proberen te schieten" *We try to shoot*) and the third allowing for both function words (i.e., "We proberen te/de schaatsen" *We try to skate* or *We try the skates*). The results showed that identification functions varied with respect to the biasing context. There were more voiceless responses in the ambiguous region of the continuum when the following word was an unambiguous verb (*schieten*) and more [də]-responses when the following word was an unambiguous noun (*schoenen*). The function for ambiguous sentences (*schaatsen*) was between the other two.

Van Alphen and McQueen (2001) also looked at the time course of the effect, as has been done in other studies (e.g., Fox, 1984; McQueen, 1991; Pitt & Samuel, 1993) by dividing the data into fast, medium, and slow time ranges. Such a split into time ranges allows for a more detailed analysis of the development of the effect over time. Van Alphen and McQueen (2001) showed that the contextual effect was strongest in the fast time range, decreased in the medium time window and vanished in the slowest responses. They proposed that the integration of syntactic information might have been completed in cases where subjects tended to give slower responses. If one assumes that syntactic information can influence phonetic decisions only as long as syntactic processing has not been completed, the absence of the sentential effect in the slow time window could be explained. In slow responses syntactic information is no longer available and has thus no influence on phonetic decisions any more. Van Alphen and McQueen (2001) investigated in a second experiment whether sentential effects would be absent when disambiguating information had not yet been provided. They constructed context sentences in which disambiguating information was delayed for longer than in Experiment 1. The ambiguous sentence from the previous experiment ("We proberen te/de schaatsen ... ") was extended by three different types of prepositional phrases. With that procedure they again created a [tə]-biased context (e.g., "We proberen te schaatsen op noren" *We try to skate on racing skates*), a [də]-biased context ("We proberen de schaatsen van mijn broer" *We try the skates of my brother*) and an ambiguous context ("We proberen te/de schaatsen zonder sokken" *We try to/the skate(s) without socks*).

If the hypothesis were true that syntactic information can exert its influence on phonetic categorization only in a restricted time window, then the overall pattern should be the same as in the previous experiment while the time pattern should be reversed. Thus, fast responses should tend to be initiated before disambiguating information arrived, so that no shift in the categorization function in the fast time window should be expected. Although the experiment failed to produce stable effects, there was a trend in the expected direction. There was no context effect in the fast and medium time ranges but an almost significant effect in the slow RT window. Van Alphen and McQueen (2001) took the combined results from these experiments as evidence for a time-limit on the effectiveness of syntactic bias on phonetic categorization. As soon as the processing of syntactic information has been completed, there is no way in which this information can influence phonemic decisions any further.

To summarize, the studies discussed so far have convincingly demonstrated that sentential context - either syntactic or semantic in nature - biases listeners' identification of ambiguous sounds. This is true for sounds that are part of both content and function words. Furthermore it seems that these biases are subject to tight temporal constraints. They can only be effective if the underlying processes have not been completed before the listeners' responses are made or if those processes have already been initiated.

If there is a sentential influence of syntax on the identification of function words, there could also be similar influences on the identification of inflectional morphemes. As already mentioned, these elements are *syntactically* predictable, while phonemes as parts of word stems are not. But in order to attribute such potentially different effects for inflected verbs vs. uninflected nouns to the processing of the context, it would be necessary to also look at the relevant target words in isolation and to demonstrate that the differential effect vanishes. If this differential effect, however, does not vanish when verbs and nouns are presented in isolation, one would have to conclude that there are more basic processing differences between the two categories. This would imply that *inflected verbs* and *uninflected nouns* are accessed differently irrespective of the sentential contexts they are produced in. What should be obtained in any case is a lexicality effect for both inflected verbs and uninflected nouns when they are presented in isolation. The interesting question, therefore, is whether the categorization functions for isolated verbs and nouns would look different from functions for verbs and nouns in appropriate sentence contexts.

Lexical effects

So far, lexical effects have only been reported for nouns and *uninflected* verbs but not for *inflected* verbs. The first study that showed involvement of the lexicon in phoneme identification was that reported by Ganong (1980). He created matched word-nonword and nonword-word voicing continua where in one group the voiced endpoints formed words whereas the unvoiced did not, while in a second group the unvoiced endpoints formed words and the voiced did not (e.g., [dæʃ] vs. [tæʃ] and [dæsk] vs. [tæsk], respectively). Subjects in this study gave more voiced responses in the ambiguous continuum region when the voiced endpoint formed a word (i.e., *dash*) but gave more voiceless responses when the unvoiced endpoint formed a word (i.e, *task*). This so called lexical effect in phonetic decision-making has been replicated many times (e.g., Connine & Clifton, 1987; Fox, 1984; Miller & Dexter, 1988; Pitt, 1995; Pitt & Samuel, 1993). It has also been demonstrated that lexical effects can be obtained when subjects are asked to label word-medial phonemes (e.g., Connine, 1990). Connine (1990) constructed stimuli in which the medial stop consonant was manipulated on a place of articulation continuum. In one condition the velar endpoint of the continuum formed a word (e.g., *bagel*) while in the second condition the alveolar endpoint was a word (e.g., *cradle*). Not surprisingly, the results showed that lexical status influenced listeners' performance such that more [d]-responses were observed when the [d]-endpoint was a word (i.e., *cradle*) but more [g]-responses were given when the [g]-endpoint formed the word (i.e., *bagel*).

Of greater importance for the current study are results from a study that showed lexical influences on word-final phoneme categorization conducted by McQueen (1991, see also Pitt & Samuel, 1993). Because inflectional morphemes in Dutch occur at the ends of words (with one exception, namely the prefix *ge-* in Dutch participles like *gestolen* 'stolen'), a study on the influences of sentential context on the processing of morphemes should look at word-final sounds. In McQueen's study listeners were asked to judge whether the last phoneme of a presented word was an [s] or an [ʃ]. Three different [s]-[ʃ] continua were tested: in one condition the [s]-endpoint formed a word (e.g., *kiss*) while in the other condition the [ʃ]-endpoint was a word (e.g., *fish*) and in a third condition both endpoints formed nonwords (e.g., *jish* and *jiss*). In an overall analysis there was no shift in the categorization functions as a reflection of a lexical effect. When McQueen (1991) split the data into fast, medium and slow RT ranges, he only found an inverse lexical effect in the medium time window. There were more [ʃ]-responses in the ambiguous region for the condition where the [s]-endpoint formed the word. Following Connine and Clifton (1987), McQueen also conducted an RT analysis in which the RTs obtained for lexically consistent and inconsistent endpoint stimuli were compared. In contrast to the Connine and Clifton (1987) results (i.e., those obtained with word-initial sounds) he found that word responses at the endpoints of continua were reliably faster than nonword responses.

But why was there no lexical effect in the categorization functions? McQueen hypothesized that the material were of too high quality, so that subjects were able to perform the task without the assistance of the lexicon. He thus ran the same experiment again with degraded stimuli (degradation was obtained by low-pass filtering the material with a cutoff of 3 kHz). This time there was a lexical effect, which was confined to the ambiguous region of the continuum. Inspection of fast, medium and slow responses showed that this lexical effect was attributable to listeners' faster reactions. McQueen (1991) argued that - because of the high quality of the stimuli - listeners in the first experiment were able to solve the task on the basis of the acoustic-phonetic information alone and could thus ignore lexical information. Only when the material was degraded did listeners start to use the higher-order information to perform their task. However, Pitt and Samuel (1993) demonstrated that degradation of stimuli is not necessarily a prerequisite for lexical effects on the identification of word-final ambiguous phonemes. On a [b]-[m] continuum, stimuli varied from the word [krɪb] to the nonword [krɪm] and from the nonword [swɪb] to the word [swɪm]. On a second [g]-[k] continuum, [rʌg] - [rʌk] and [stʌg] - [stʌk] formed the stimulus pairs. The quality of both continua was manipulated so that there was a clear version and a noise version of each continuum. In both conditions, Pitt and Samuel (1993) found reliable function shifts towards the nonword endpoints of the continua. And although the shifts in the categorization functions were larger in the noise condition than in the clear condition, this difference hardly ever reached significance. Thus, phonetic categorization on word-final ambiguous sounds can reveal lexical effects also with high-quality material.

It is interesting to note, however, that the occurrence of lexical effects is to some extent dependent on the careful construction of the experimental stimuli or on explicitly manipulated task demands. For example, Eimas, Hornstein, and Payton (1990), among others, demonstrated that the occurrence of lexical effects during phoneme monitoring could be manipulated by shifting the listeners' attention from the prelexical to the lexical processing level. In the phoneme monitoring task, listeners are usually asked to monitor auditorily presented words or nonwords for phoneme targets defined before the experiment. Classically, listeners can perform this task faster when the target phoneme is presented in an existing word as compared to a target-bearing nonword (see, for more details, Connine & Titone, 1997; Cutler, Mehler, Norris, & Seguí, 1987; Pitt & Samuel, 1995; Frauenfelder et al., 1990). Eimas et al. (1990) used the phoneme monitoring task in a slightly adapted version: they asked subjects to decide which of two prespecified phonemes constituted the initial sound of a target word presented at the end of a short neutral carrier phrase ("The next word is ..."). On a word like *band*, for example, they had to decide whether the first phoneme was a voiced [b] or a voiceless [p]. Eimas et al. (1990) only observed an RT advantage in the monitoring latencies for target-bearing words as compared to nonwords when subjects were asked to perform a secondary task that involved, for example, a lexicality judgment (i.e., whether the relevant stimulus was a word or a nonword). Interestingly, a secondary task that required non-linguistic information such as length of the given word or nonword did not reveal lexical effects. Moreover, in line with McQueen (1991), lexical effects were obtained when stimuli were degraded by pink noise. These findings were discussed within the framework of the Race model (Cutler & Norris, 1979) which will be introduced in more detail below. This model assumes that phonemic decisions might be based on either prelexical or lexical representations. Eimas et al. (1990) argued that their data suggest that listeners tend to base their phonemic decisions on prelexical information and that only additive task demands might shift the listeners attention to the (post)lexical level. It is important to note that this shift, as it were, spilled over from the secondary task (i.e., lexicality judgment) to the primary task (i.e., phonemic decision). This, following Eimas et al. (1990), might reflect the fact that listeners found it cognitively more economical to base both primary and secondary decisions on one processing level than to base the first decision on prelexical information and then shift their attention to the lexical level for the second decision.

Of particular interest was a follow-up study to these experiments by Eimas and Nygaard (1992), in which the authors tried to extend the previous findings with words presented at the end of a short carrier phrase to the same words when presented in longer and more complex sentence contexts. In several experiments, the authors failed to demonstrate equivalent lexical effects for words in sentence frames as they had previously observed for these words in the short carrier phrase. Thus, words were presented in neutral coherent sentence frames (e.g., [b] in “Congress will vote on a new BILL/BAN on foreign imports” with the target word *bill* being high-frequent and the target word *ban* being low-frequent) and listeners - after deciding about the first phoneme of the target word - had to rate the familiarity of that word. No frequency effect was observed under these conditions (see for more details Eimas & Nygaard, 1992). This was in contrast to what Eimas and Nygaard had expected: they had predicted that the presentation of target words in a more natural condition (i.e., in complex sentences rather than in a small carrier phrase like “The next word is ...”) would shift the listeners attention towards the lexical level more easily.

When the same words were predictable on the basis of the sentence context (e.g., predictable: “The doctor sent a BILL” vs. unpredictable: “The institution sent a BILL”) there were effects of predictability but there was no interaction of predictability with the factor frequency. Thus, when the target phoneme was contained in a high frequency word, this target was detected more easily when the word was predictable than when it was not predictable. The same was true for low frequency words, but these words did not elicit slower RTs than the high frequent counterparts. This was also true when a secondary task required lexical knowledge.

The only condition under which Eimas and Nygaard (1992) found substantial lexical (i.e., frequency) effects was when words were presented in contexts that consisted of unstructured word lists (e.g., “She of wonder what wants TASK/TACK”, with *task* being high-frequent and *tack* being low-frequent). Importantly, this lexical effect was only obtained when listeners had to perform a secondary task that required lexical knowledge.

Eimas and Nygaard (1992) interpreted these results as follows: when target-bearing words were presented in fully coherent structures, listeners might have found it more economical to dissociate the phonetic processing (required for the primary task of phoneme monitoring) from the sentential processing (required for the secondary task). Since the processing of the sentence context provided a fully coherent mental structure, the decision whether a target-bearing item had been a word or a nonword, for example, could be achieved by simply determining whether successful integration of that item had taken place. Only when there was no such coherent structure did listeners start to focus their attention on a single information level, namely the lexical level, to perform both the primary and the secondary tasks. Otherwise, listeners used prelexical information to perform the phoneme monitoring task while they based their lexical decisions on higher order (i.e., contextual) information. The authors therefore concluded that contextual coherence determined whether lexical knowledge was consulted in order to influence phonemic decisions.

Taken together, these data convincingly show that there are lexical influences on the way listeners identify (ambiguous) phonemes. Furthermore, it appears from the Eimas et al. (1990), Eimas and Nygaard (1992) and Miller et al. (1984) studies that these influences in sentence contexts are not guaranteed. Instead, lexical information might play a secondary role when sentential structure can provide a more economical way to solve the task at hand.

The current study uses the tool of phonetic categorization to shed light on the role of morphology in spoken language comprehension. As already outlined, the most important questions are the following: (A) is the identification of an ambiguous sound influenced differently by the sentential context when this sound is a potential inflectional morpheme (e.g.; [?] in a [t]-[k] continuum *gaat - gaak*, where *gaat* ’leaves’ is an inflected verb) as compared to when it is part of the word stem (e.g., [?] in *straat - straak*, where *straat* ’street’ is an uninflected noun) and (B) are inflected verbs and uninflected nouns treated similarly (again in the phonetic categorization task) when presented in isolation? Clearly, there are sentential as well as lexical influences on phonetic categorization; the issue to be addressed here is whether and, if so, how both these factors influence the categorization of potentially inflected forms.

Feedback or no feedback?

In contrast to most other studies in the history of phonetic categorization research, the present study does not contribute directly to one of the most hotly debated issues in comprehension modeling: the question of feedback. Nonetheless, the theoretical interpretation of the current data depends on the basic assumptions which are made about whether or not there is feedback in the speech recognition system and on where in the system phonetic decisions are made. At the core of the feedback discussion stands the following question: does a model of speech perception need a feedback device which allows for the flow of information from higher order processing levels back to lower levels in order to explain contextual effects on phonemic decisions? As discussed in more detail in Chapter 1, there are basically two accounts of human speech perception that diverge strongly with respect to the issue of feedback: autonomous (e.g., Norris et al., 2000) and interactive (e.g., McClelland & Elman, 1986) frameworks. The resolution of the feedback discussion is especially problematic since most data can be explained within both frameworks (for a detailed overview of the literature see Norris et al., 2000). Thus, the strongest arguments for or against one or the other type of model currently appear to be theoretical rather than empirical in nature. Future research, however, might be able to resolve the debate.

Both autonomous and interactive models accept the fact that lexical (or sentential) knowledge can exert an influence on perceptual decisions. They disagree, however, on *the way* in which these influences have their effect. While autonomous models assume only “bottom-up” flow of information, implying the independence of each processing stage, interactive models also allow for information to flow “top-down” and by that to alter the operation of earlier processing levels. Interactive accounts thus assume a direct influence of lexical information on early processes of phonemic analysis, whereas autonomous models assume this early analysis to be unaffected by higher-level processes. In interactive models, phonetic decisions are therefore based on the *prelexical* analysis stage (like for example the phoneme level in TRACE), while in autonomous models decision making is based on a separate process rather than on the initial phonemic analysis.

In the autonomous Race model (Cutler & Norris, 1979), for example, two sources of information are available for the identification of phonemes. There is a *lexical* route that activates phonemes via the activation of words in the lexicon and there is a *prelexical* route that activates phonemes on the basis of the acoustic-phonetic input alone. In the Race model, lexical effects in phonetic categorization, for example, are the result of a race between these two routes. According to the Race model, responses to ambiguous phonemes will sometimes be based on the lexical route and sometimes on the phonemic route, and the bias in the categorization function is the result of the contribution of the lexical route. Note that the Race model assumes direct activation of lexical elements in the lexicon via which certain phonemic decisions can be biased. The prelexical percept however remains unaltered by that process.

Because lexical effects in the Race model depend on successful lexical access, studies that demonstrate lexical effects on nonwords that resemble existing words challenge the Race model (Connine, Titone, Deelman, & Blasko, 1997; Marslen-Wilson & Warren, 1994; McQueen, Norris, & Cutler, 1999; Newman, Sawusch, & Luce, 1997; Wurm & Samuel, 1997). For example, instead of presenting listeners with the classical word-nonword continua in phonetic categorization, Newman et al. (1997) created pairs of nonwords in which they manipulated the neighborhood density of the nonwords so that in one series the voiced endpoint was more “word-like” (i.e., higher neighborhood density) while the reverse was true in another series (i.e., higher neighborhood density for the voiceless endpoint). Their argument was that not only the lexical status of a given stimulus should influence listeners’ decisions about ambiguous phonemes but also the set of words similar in sound to a presented stimulus. Thus, the more word-like a given nonword is, the higher the respective neighborhood density is, and thus more responses in an ambiguous region should be consistent with that nonword. Newman et al. (1997) indeed found that there were more voiced responses when the voiced endpoint had more lexical neighbors than the voiceless endpoint (e.g., [gɑɪs] vs. [kɑɪs]). The opposite was true when the voiceless endpoint nonword had a higher neighborhood density (i.e., [gɑɪp] vs. [kɑɪp]).

Similarly, Connine et al. (1997) demonstrated lexical effects on the processing of nonwords with the phoneme monitoring task. Connine et al. (1997) found increasing RTs for phoneme monitoring in nonwords as a correlate of decreasing similarity of the respective nonwords to real words. Nonwords were constructed by changing the initial phonemes of existing words according to one phonetic feature (e.g., *gabinet* from *cabinet*) or six phonetic features on average (e.g., *mabinet* from *cabinet*). The less phonetic features were altered, the more the resulting nonwords resembled real words, and the faster were the RTs. The Race model cannot account for these results because there are no entries for word-like nonwords in the lexicon that could be involved in a race between lexical and phonemic information. An interactive account, on the other hand, can easily explain the results in the following way: because phonemes in word-like nonwords receive feedback from those words that they have partially activated, their activation in turn is boosted more quickly and more strongly than the activation of phonemes that are parts of nonwords less similar to existing words.

Are lexical influences on the processing of nonwords therefore strong evidence against an autonomous approach to spoken language comprehension? The answer is no. Although the Race model cannot account for the data described above, another autonomous model, namely Merge (Norris et al., 2000), can explain these results without the need for feedback. In Merge, phonemic and lexical information can join forces to determine phonemic identification responses. This is achieved by assuming phonemic decision units where the activation of representations at both the phonemic and the lexical level is integrated. Lexical effects in nonwords are thus explained by the partial activation of similar sounding words whose activation in turn biases the activation of the relevant decision units. Note that this flow of information from the (lexical) level of form representations to the (postlexical) level of decision units is unidirectional in nature.

Other studies have been conducted, however, that on first sight challenged the autonomous view. Of particular importance for the feedback debate was a study conducted by Elman and McClelland (1988). Based on an earlier study by Mann and Repp (1981), Elman and McClelland demonstrated that listeners compensated for coarticulation when asked to label stop consonants ambiguous on a place of articulation continuum between [t] and [k] that were preceded by either the palatal fricative [ʃ] or the alveolar fricative [s]. In speech production the articulation of the velar [k], for example, will be more anterior when preceded by the alveolar [s] then when preceded by [ʃ]. The place of articulation of the fricative thus influences the articulation of the following stop consonant. Listeners compensate for this coarticulation when asked to label sounds that are ambiguous on a place of articulation continuum between [t] and [k] depending on the preceding fricative. An ambiguous sound that lies more towards the [t] endpoint of the continuum will still be labeled as [k] when preceded by the fricative [s]. Listeners therefore seem to be ’aware’ of the fact that a velar stop after [s] sounds more alveolar than in a neutral context and therefore tend to accept ambiguous sounds on the [t]-[k] continuum in this context as [k]. The opposite shift of the category boundary is observed for ambiguous stop consonants that follow the palatal [ʃ].

Elman and McClelland (1988) obtained a similar shift in the categorization function when listeners were asked to label a word-initial ambiguous sound as either [t] or [k]. These words were preceded by fricative-final words like *christmas* or *foolish*. Just as in the Mann and Repp (1981) study, there were more [t] responses in the ambiguous region of a *tape-cape* continuum after words like *foolish* and more [k] responses after words like *christmas*. But more important for the feedback discussion was an additional manipulation: the final fricatives of words like *christmas* or *foolish* were replaced by an ambiguous sound that was midway between [s] and [ʃ]. Listeners still showed a compensation for coarticulation effect. This effect, as Elman and McClelland pointed out, must have been mediated by the lexicon. Because an ambiguous fricative [?] will be perceived more like an [s] in a *christma?* surrounding (the classical lexical effect), listeners gave more [k] responses after these words, just as they gave more [t] responses after the ambiguous fricative-final word *fooli?*, where [?] will be perceived more as an [ʃ]. Thus listeners behaved as if they had heard unambiguous fricatives instead of ambiguous ones. As already outlined above, the process of compensatory perception in this study (and the previous Mann & Repp, 1981 study) is based on the fricative information of words like *christmas* and *foolish*. Whether this information,according to Elman and McClelland (1988), is supplied by the acoustic signal itself “bottom-up” (as in the unambiguous version) or by the lexicon “top-down” (as in the ambiguous version) is not important for the triggering of the compensation process. The strong point in favor of interactive models was that the lexicon was able to influence a process - namely that of compensation for coarticulation - which is supposed to occur prelexically.

In defence of the autonomous view, however, some authors have argued that these effects were not necessarily the result of a direct influence of the lexicon but rather were the result of a sensitivity to the transitional probabilities between phonemes (which acts at the prelexical level). The argument goes as follows: all [s]-final words (e.g., *christmas*) in the Elman and McClelland study ended with the VC cluster [əs] while all [ʃ]-final words (e.g., *foolish*) ended with the VC string [ɪʃ]. The final fricatives were therefore predictable on grounds of sequential probabilities in the English vocabulary: the sound [s] is more likely to occur after [ə] whereas [ʃ] is more likely to occur after [ɪ]. The lexicon is thus not the only potential mediator of the compensation effects. The perception of an ambiguous fricative at the end of the string *fooli?* could be biased towards the fricative [ʃ] on the basis of the preceding vowel rather than on the basis of lexical information. It has also been demonstrated by Norris (1993) and by Cairns, Shillcock, Charter, and Levy (1995) that a recurrent network without lexical knowledge can produce compensation effects on the grounds of bottom-up information alone (for more detailed information see also Pitt & McQueen, 1998 and Norris et al., 2000).

Empirical evidence for the autonomous explanation of the Elman and Mc-

Clelland results has more recently been provided by Pitt and McQueen (1998). These authors disentangled the two confounding factors of lexical information and transitional probabilities. In a series of experiments, listeners were asked (just as in the Elman and McClelland study) to label word-initial stop consonants that were ambiguous between [t] (*tapes*) and [k] (*capes*). In addition, listeners were required to judge the preceding fricatives as either [s] or [ʃ]. Pitt and McQueen (1998) distinguished the following two bias conditions: (A) there were either lexical biases towards one or the other fricative (e.g., *juice* and *bush*) while the transitional probabilities (TPs) were kept constant (both [u] and [ʊ] have equal TP biases towards [s] and [ʃ]), or (B) there were TP biases without lexical involvement (the nonword sequence *der* [dɝ] has sequential biases towards [s] while the nonword sequence *nai* [neɪ] has such biases towards [ʃ]). These different biasing contexts could either be followed by unambiguous instances of [s] and [ʃ] or by an ambiguous sound midway between the two. These three different fricative sounds were presented in either lexical or TP contexts and all of those contexts were followed by words whose initial sounds varied on an eight-step place of articulation continuum.

In the **unambiguous** fricative cases, Pitt and McQueen (1998) reported a compensation for coarticulation effect for both lexical and TP contexts: there were more [t] responses after an unambiguous [ʃ] and accordingly more [k]-responses after an unambiguous [s]. No lexical or TP influences were observed when the fricatives were unambiguous. Thus, the **identity** of the fricatives was the determining factor rather than **fricative bias** (which could have been mediated by both lexicality or TP). Because the acoustic signal was clear enough, no higher-order information exerted any influence on listeners’ perception.

In the **ambiguous** fricative condition, however, Pitt and McQueen (1998) found a compensation for coarticulation effect in the TP condition but not in the lexical condition. Remember that the transitional probabilities in the lexical condition were kept constant across the items *juice* and *bush*. On the other hand, transitional probabilities in the nonword condition were manipulated (so that one VC sequence favored [s] and the other favored [ʃ]). And indeed Pitt and McQueen found a stable compensation for coarticulation effect in this condition. Thus, there were more [t] responses following the nonword *nai?* which had TP biases towards [ʃ]: the ambiguous fricative was thus perceived by listeners as if it had been an unambiguous [ʃ] which in turn triggered the compensation effect. Accordingly, there were more [k]-responses after the nonword *der?*, implying that listeners perceived the ambiguous fricative more like an [s].

If compensation for coarticulation following ambiguous fricatives was mediated by the lexicon, there should have been such an effect in the *jui?-bu?* condition. The lexicon indeed biased listeners' perception of the ambiguous final phonemes towards [s] in the *jui?* but towards [ʃ] in the *bu?* case as shown by the listeners' fricative judgements. In an interactive model, this lexical involvement should have triggered the compensation effect. This is not what Pitt and McQueen observed.

Even if Pitt and McQueen (1998) have neatly demonstrated that the Elman and McClelland (1988) study was not as strong an argument against autonomy as some authors had assumed, the debate about feedback continues to be as lively as ever. Samuel (reported in Samuel, 2001), for example, have recently demonstrated a compensation for coarticulation effect with ambiguous fricatives in lexical contexts like those of Elman and McClelland (1988) but with the transitional probabilities of the fricatives controlled.

Another study challenging autonomous models (probably the most challenging one so far) is Samuel (1997). In this study, the two effects of phoneme restoration (as had been established by Warren back in the 1970s) and selective adaptation (e.g., Eimas & Corbit, 1973; Sawusch & Jusczyk, 1981) were combined. In adaptation experiments it has been demonstrated that listeners' identification of sounds that vary along a certain phonemic continuum can be manipulated by previous adaptation of the listeners to one or the other endpoint phoneme. Thus, when listeners are presented repeatedly with a sound sequence like for example [gɪ] (as in *gift*), they are less likely to label a sound ambiguous between [g] and [k] as voiced than before the adaptation conditions. Crucially, the categorization functions for the same ambiguous items differ as a direct reflection of previous exposure to adaptation.

Samuel (1997) demonstrated that this adaptation effect also occurred when phonemes in adaptor words had been replaced by noise. Subjects were asked to identify a sound sequence as either [bɪ] or [dɪ] that varied on an eight-step place of articulation continuum. Before the categorization part they were adapted to one of the two endpoint phonemes [b] or [d] through exposure to real English words that contained either one of the critical sounds (e.g., *confidential* or *exhibition*). As expected, listeners' categorization functions varied according to the previous exposure conditions. When Samuel replaced the [b] and [d] sounds in his adaptors with white noise (i.e., *confi∗ential* and *exhi∗ition*) a similar adaptation effect on listeners' identification performance was obtained. However, when the critical phonemes had been replaced by silence (i.e., *confi_ential* and *exhi_ition*) there was no adaptation effect any more.

Samuel (1997) therefore concluded that the white noise in his adaptors was not only restored by the listeners but moreover produced an adaptation effect just as if the 'real' phonemes had been present. In other words, the lexicon filled in the missing acoustic information so that the physical signal was replaced by a lexically-derived percept. For listeners' identification performance it did not matter whether they had heard a real [d] or [b] or whether these phonemes had been restored by the lexicon. This conclusion was supported by the failure to find an adaptation effect when the [b] and [d] sounds within the adaptors had been replaced by silence. Samuel (1997) argued that the activation of lexical items (like *confidential* and *exhibition*) can directly influence the perceptual process such that phonemic information is created online.

No direct empirical evidence has yet been produced that disproves Samuel's claims. But there has of course been determined protest against Samuel's conclusions that his data strongly support the idea of a feedback device in the perception system. One important argument against Samuel's conclusion is that it is far from clear at which stage of processing adaptation happens. As Norris et al. (2000) emphasize, it has been demonstrated that adaptation can occur at different levels of the processing system. Norris et al. (2000) suggest that the locus of the adaptation effect may be at the decision nodes in Merge. If so, than the lexical influence on adaptation observed by Samuel (1997) can be explained without feedback.

But why is it so important to avoid feedback in the perceptual system? The main advantage of unaltered perception of acoustic reality - as assumed by autonomous models - is that listeners are protected from hallucinating. While *perceptual illusions* are a well-documented phenomenon not only in the psycholinguistic literature (e.g., phonemic restoration) that might even improve the quality of the perceptual input, *hallucinations*, at least in healthy people, do not occur. For example, as has been described above, Samuel (1997) found no adaptation effect when the phonemes in his adaptor words had been replaced by silence. Thus, when there was no bottom-up information of any kind (not even noise), feedback alone did not create information from nothing. The question thus arises why one should build in a mechanism that could - theoretically speaking - result in hallucinations although - empirically speaking - these hallucinations do not occur? In this sense, feedback is a bad design feature of a model of speech perception, because it allows for a phenomenon that is not evident in the cognitive performance of healthy people. Furthermore, as Norris et al. (2000) argue, feedback cannot even improve the word recognition process (see for more detailed arguments Norris et al., 2000). And because feedback cannot assist the word recognition process, interactive models like TRACE violate the principle of parsimony, that models should be as simple as possible (Occam's razor, see Norris et al., 2000).

Empirical data that unambiguously rules out one or the other type of model has not yet been provided. Thus - even if theoretical arguments favor the autonomous view - the question about feedback in the human speech recognition system will undoubtedly remain a topic that will exercise the brains of researchers in the future.

Modeling morphological effects

As already mentioned, the current study was not designed to contribute directly to the debate about feedback. Still, it needs to be clarified in what way morphological effects in phonetic categorization might be explained within one or the other model. Remember the core questions of the present study: is the identification of an ambiguous sound influenced differently by the sentential context when this sound is a potential inflectional morpheme (e.g., [?] on a *de tante gaat - de tante gaak* continuum, with ’gaat’ as the inflected verb) than when this sound is part of the word stem (e.g., [?] on a *een brede straat - een brede straak* continuum with ’straat’ as the uninflected noun)? And if there is such a differential effect, would this difference also hold for inflected verbs and uninflected nouns presented in isolation? How then would sentential or lexical effects on inflectional morphemes fit into current models of speech recognition?

Morpho-syntactic effects in Merge and TRACE

As van Alphen and McQueen (2001) demonstrated, the sentence level in an autonomous model like Merge can have a direct influence on the identification of ambiguous function words. In Merge, identification is mediated by dedicated decision units set up at a postlexical decision-making stage. Because the inflectional morpheme ’-t’ in the phrase ’de tante gaat’ is predictable on the basis of the *syntax*, it is plausible to assume that the sentential level - in addition to the lexical level - can bias the decision-making stage towards [t]-decisions when the model is confronted with an ambiguous sound. If the final sound of a sentence was perfectly ambiguous, decisions in the VP case could be potentially affected by three sources of information (phonemic, lexical, and sentential) while decisions in the NP condition (*een brede straat*) should only be influenced by the phonemic and the lexical level. There would thus be a stronger bias towards ’-t’ in a sentence like ’de tante gaat’ as compared to the noun phrase ’een brede straat’. Although sentential context could perhaps be created such that it imposed strong constraints on the semantic appropriateness of content words like ’straat’, the sentences used in the current study were not of that kind. The sentence frame ’een brede’ is not highly restrictive in what nouns should follow. It certainly does not require a noun that ends with a ’t’. Similarly, the VP ’de tante’ is semantically not highly predictive of the verb that should follow. However, it is more constrained in a *syntactic* sense: if a verb follows the noun ’tante’ this verb needs the inflectional marker ’-t’ in order to form a grammatical sentence.

In order to allow for sentential information to modulate a phonemic decision in Merge, one extra assumption is necessary. Generally, the sentence level is supposed to be a conceptual processing stage that is free from form information. Once a lexical entry has been accessed via its form representation, this form information should not be relevant anymore. But in order to allow such a conceptual level to have an influence on a form-based level like the decision-making level in Merge, some form information must be accessible at that stage. Thus, either this information is stored centrally or it is retrieved, as it were, on its way to the decision level. Because the decision level in Merge is an output level where reactions are initiated, one might assume that the activation which is passed forward from the conceptual level to the decision level might be mediated via the speech production system (which is an output-oriented system just like the decision stage in Merge). The necessary form information could thus be retrieved rather than stored at the central level.

There is of course another possibility to explain potential sentential effects on inflectional morphemes: feedback as implemented in interactive models like TRACE. In TRACE, sentential information - just like lexical information - could be fed back to lower processing levels and thus bias the identification of ambiguous sounds. In a VP that required the inflectional marker ’-t’ on its final word, a phonemic decision would thus receive feedback activation from two higher-order processing levels (namely the sentence level and the lexical level). Because semantic predictability was less constrained than syntactic predictability, less feedback activation should bias the phonemic categorization of the ambiguous sentence-final sound in the noun ’straa?’ as compared to the same sound in the verb ’gaa?’.

Access representations

In both models, sentential effects might be mediated either via access representations for inflectional morphemes or via full-form representations of inflected words. It is thus important to note that a sentential effect that was different for VPs as compared to NPs would not allow for any conclusions about the structure of the lexical level. What can shed light on the debate about decomposed or full-form representations of morphologically complex words, however, is a comparison of the categorization functions for inflected verbs and uninflected nouns presented in *isolation*. It is possible that there is a stronger bias towards [t] when the ambiguous sound is part of an inflected verb than when the same sound is part of an uninflected noun. One way to account for such a difference in categorization behaviour would be to assume separate access representations for inflectional morphemes. In the present study, the activated access representation of the stem ’ga-’ could feed activation to a linked access representation of the inflectional morpheme ’-t’ (also represented at the lexical level). In Merge, this activation would in turn activate the decision node for [t] at the decision stage and would thus result in a bias towards more [t] responses in the ambiguous region of the continuum. Similarly, in TRACE, the activated stem ’ga-’ could also activate the morpheme ’-t’ and both would feed back information down to the phoneme level to bias the phonemic decision towards [t].

But note that this effect is no different from a classical *lexical* effect (which should also be observed for the noun ’straat’). In order to explain a stronger ’lexical’ effect (a morphological effect) for isolated inflected verbs than for isolated uninflected nouns, one might assume that the link between the inflectional morpheme ’-t’ to the [t]-decision unit is in some sense a special one. It is possible that inflectional morphemes have stronger connections to decision units because of their special syntactic status. Whatever the appropriate implementation of a morphological effect in the categorization of isolated words might be, such an effect would suggest that morphological structure was represented in some way in the mental lexicon. This, moreover, would be consistent with a model in which there were decomposed access representations. One of the key questions in this study is therefore whether there is a difference in categorization performance between words presented in isolation as compared to words presented in sentential context.

The feedback issue

So far, none of the potential results outlined above would allow for a differentiation between autonomous and interactive models of spoken language comprehension. The time course of the categorization effects might however be informative in this respect. Because TRACE assumes that activation from higher-order levels on the perceptual level accumulate over time, sentential as well as lexical effects in TRACE should be quite stable over time or, if anything, should increase over time. Thus, TRACE would predict similar patterns over time for both the verbal and the nominal conditions. This is true although TRACE has the built-in time constraint that guarantees that any influence of sentential information will be effective only in conditions where contextual information has had the opportunity to feed back via the lexical level to the perceptual level. Because categorization in the current experiment is required on the very last phoneme of each utterance, one might expect a lexical effect already in the fast responses but one should definitely expect such an effect in the slower reactions. Note also that TRACE assumes the perceptual decision units to be the same as the encoding units. Once such a unit has been altered by higher-order information (either sentential or lexical) this alternation cannot be undone. Thus, listeners should also show a bias towards the [t] endpoint in both the verbal (*de tante gaat*) and the nominal (*een brede straat*) conditions in their slower responses.

Merge, however, predicts a decrease of lexical as well as of sentential influence on phonetic categorization over time. In both the verbal and the nominal phrases, the target sound is the final sound of the whole phrase. Merge assumes that once lexical and/or sentential information has been resolved (i.e., integrated into the interpretation of the utterance), this information loses its influence on phonetic decisions (see, for example, van Alphen & McQueen, 2001). In fast responses, when the relevant information has not yet been fully integrated, the higher-order influences should thus be strongest. This influence should vanish over time. While neither of the models predicts a difference between the time patterns of the categorization functions for NPs as compared to VPs, the two models do predict different overall time patterns.

Weaker effects for inflected verbs?

So far it has only been hypothesised that - if there was a different effect for inflected verbs than for uninflected nouns in sentences - the processing of syntactic information should result in a *stronger* sentential/lexical effect for inflected verbs than for uninflected nouns. There is, however, one other possibility: if there is a sentential/lexical effect it might be stronger for uninflected nouns than for inflected verbs (both presented in sentence contexts) or this potential effect might even be absent for inflected verbs. This might be due to morphological parsing: if the ambiguous sound was in a morpheme position (as is the case in the verbal phrase ’de tante gaa?’) this sound could be more easily separated from the rest of the sentence, and thus might be categorized, as it were, independent of the context. That is, if an ambiguous sound occurs in a position where it would perform a morpho-syntactic function it could thus be assigned an independent status in the ongoing parse of the sentence and thus could be treated, in the phonetic decision, as an entity separate from its context. More generally, this would suggest that, given a sufficiently constraining context, listeners can dissociate the process of syntactic integration from that of phonetic decision-making (see, for example, Eimas & Nygaard, 1992). Such a dissociation seems less likely for nominal sentence contexts since the final phonemes of the relevant nouns are not decomposable morphological units.

This leads to the following possibilities: the dissociation of the two tasks -

namely, morphosyntactic parsing and integration, on the one hand, and phonetic decision-making, on the other hand - might only work well in an appropriate sentence context. That is, it might work for inflected verbs in sentence contexts but not for uninflected nouns in sentence contexts. Furthermore, such a dissociation might only work when there is sufficient syntactic context that licenses the decomposition of the final verb. If so, one might not expect this dissociation to work on inflected verbs presented in isolation because there is no preceding context that would require and/or license the decomposition of the inflected verb. However, it might be the case that listeners used their lexical knowledge - instead of higher-order syntactic information - when inflected verbs are presented in isolation (see, for example, Eimas et al., 1990). Thus, while there might be no lexical effect for inflected verbs in sentence contexts (due to a dissociation of the tasks), such an effect might be expected for the same verbs in isolation. Since there is no sentential context that licenses the decomposition of inflected verbs, however, such a lexical effect for inflected verbs in isolation might be weaker than for uninflected nouns in isolation. That is, the lack of licensing might make an isolated inflected verb less “word-like”. Note that these potential differences between inflected verbs and uninflected nouns both in sentential contexts and in isolation suggest that morphological information is represented in the mental lexicon.

Experiment 4.1: Phonetic categorization - a pilot study

To summarize, the present study sought a better understanding of how and when morphological information might be integrated into sentences, and of how far the processing of morphological information depends on the processing of an appropriate syntactic structure. Therefore, potential morphological effects for inflected verbs are compared to lexical effects for uninflected nouns, both in sentence contexts and in isolation. The results not only have interesting implications for models of morphological processing but they might also shed light on the feedback-debate addressed by interactive and autonomous models. A series of seven phonetic categorization experiments was conducted to answer the questions outlined above. In addition, the time course of potential effects was inspected to serve two main goals: one concerns potentially different categorization patterns for inflected verbs as compared to uninflected nouns, while the other concerns the issue of whether feedback has a role to play in speech perception or not. Remember that autonomous models predict syntactic effects to be strongest in the listeners’ fastest responses, due to a limited time window in which the relevant information is available, while interactive models predict stronger effects in the slower responses since sentential effects build up over time in these models.

In the first experiment, VPs and NPs were contrasted whose final phonemes varied on a place-of-articulation continuum between [t] and [k]. In the verbal condition, the [t] endpoint of the continuum was an inflectional morpheme of the final word (e.g., [t] in *de tante gaat - gaak*) while the same sound in the nominal condition was part of the noun stem of the final word (e.g., [t] in *een brede straat - straak*). In Dutch there are no other inflectional morphemes consisting of a single stop consonant (such as the [t]). Thus, the [t] endpoint was always the endpoint that formed a real word, while the other endpoint by definition formed a nonword (at least in the VP condition). If one endpoint always forms a word, a possible morphological effect on phoneme identification (i.e., more [t] responses in the ambiguous area for verb phrases as compared to noun phrases) might be masked by a lexicality effect. Therefore it was important to also include a condition in which both endpoints formed nonwords, that is, a condition where no lexical shift should be expected. These nonword-nonword continua (e.g., *klaat - klaak*) were presented in both the verbal (*de tante klaat/klaak*) and the nominal contexts (*een brede klaat/klaak*). Thus, the critical continua could not be compared with each other but instead were judged relative to the respective nonword-nonword continua.

In order to investigate the issue of morphological decomposition further, another condition was tested in the first experiment. Nonwords with real words embedded within them (e.g., ’vla’ *custard* + [t] = ’vlaat’) were presented in final position of both the verbal and the nominal surroundings (e.g., *de tante vlaat* and *een brede vlaat*). Would listeners be more willing to label an ambiguous sound as [t] when, on the basis of the context (*de tante vlaat*, i.e., *the aunt custard-s*), they could decompose the last nonword into two meaningful units than when this decomposition would not make any sense (*een brede vlaat*, i.e., *a wide custards*)?

It is important to note that the current study differs in some respects from earlier categorization experiments. First, word-final phonemes have previously only been studied when presented in isolated words, never when presented in a sentential context. Second, the phonemes under investigation have so far always been parts of word stems - either function or content words. Single phoneme morphemes have never been looked at in phonetic categorization. And finally, categorization continua always consisted of word-nonword and nonword-word stimuli (e.g.,*fish-fiss* and *kish-kiss*; McQueen, 1991) while in the present study word-nonword (e.g., *straat-straak*) continua were compared to nonword-nonword continua (e.g., *klaat-klaak*) in order to establish lexical, sentential or even morphological effects. Because of these new manipulations, Experiment

4.1 was conducted as a pilot study in order to see whether these manipulations would produce interpretable effects.

Method

Materials. Two different types of preceding sentential context were constructed. One type ended with a verb (VP) while the other ended with a noun (NP). For each type of sentential context two different variations were constructed which yielded a total of four different context sentences. One set of contexts ended with monosyllabic words while the other set ended with bisyllabic words. The mean surface frequency for the real verbs was 7533 (per 42 million words based on the CELEX computerised database of Dutch) while the mean surface frequency of the nouns was 3136 (per 42 million words). There were four final items in each set: a verb, a noun, a pseudoword, and a nonword. These four items were matched phonologically such that only the onsets differed while the final vowels and consonants were kept constant (e.g., *gaat, straat, vlaat, and klaat*). All words were chosen so that they all formed nonwords when the final sound was replaced with a [k]. Each [t]-final item was thus paired with a [k]-final nonword to form four different continua within each set: verb-nonword (e.g., *gaat - gaak*); noun-nonword (e.g., *straat - straak*); pseudoword-nonword (e.g., *vlaat - vlaak*); and nonword-nonword (e.g., *klaat - klaak*). Each pseudoword-nonword and each nonword-nonword continuum was presented in both the verbal and the nominal contexts, while the verb-nonword and noun-nonword continua were only presented in the syntactically appropriate contexts. A full list of the 12 experimental sentences is given in Appendix C.

Stimulus construction. All 12 phrases were recorded by a female native speaker of Dutch in a sound-attenuated booth. The stimuli were recorded onto DAT tape (sampling at 48 kHz, with 16-bit resolution) and were digitized afterwards onto a computer at a sampling rate of 16 kHz. These tokens were edited with the Xwaves/ESPS waveform editor. A fifteen-step place-of-articulation continuum was created. A [t] and a [k] were spliced out of natural waveforms taken from the verb *gaat* and the matched nonword *gaak*, with the cuts being made at zero-crossings. These phonemes were spliced such that they were of equal length (212 ms). Thirteen intermediate stimuli were constructed using a procedure developed by Stevenson (1979) and Repp (1981). The amplitudes of the two waveforms were added sample by sample in different proportions. The vowel preceding the last sound was kept constant within each item set. In order to keep the transitions in the preceding vowel constant so that any bias towards [t] or [k] would be the same for each word or nonword, the same vowels were used in all contexts (within each item set). The vowels were taken from velar contexts (e.g., *gaak*), so had transitions consistent with a [k]. Any bias based on these transitions would therefore work against the predictions. The fifteen tokens from the continuum were then spliced onto all contexts.

In order to judge which tokens of the continuum were the most ambiguous ones, a pretest was conducted. Eight native Dutch speakers were asked to label the sentence-final sounds as either [t] or [k]. They were presented with a subset of the experimental stimuli: all verb-nonword (e.g., *gaat-gaak*) and noun-nonword (e.g., *straat-straak*) continua were presented in their appropriate contexts, while the pseudoword-nonword (e.g., *vlaat-vlaak*) and nonword-nonword (e.g., *klaat-klaak*) continua were only presented in one of the two contexts they could occur in. For one item set the pseudo-nonword continuum was presented in the verbal context, while the nonword-nonword continuum was presented in the nominal context. For the other item set the opposite was true so that the types of contexts that were presented were balanced within the pretest. Each of the 15 steps in each of the eight continua was presented once to the listeners. There was therefore a total of 120 presentations per subject. On the basis of visual inspection of the resulting categorization functions, steps 4 - 11 were chosen for presentation in the experiment. Steps 1 - 4 were uniformly labeled as [t], while steps 11 - 15 were uniformly labeled as [k]. Steps 4 and 11 of the continuum were thus clear instantiations of the respective sounds while the six intermediate steps were ambiguous. Henceforth the steps will be numbered 1 - 8.

Subjects. Twenty students (13 female and 7 male) were recruited from the subject pool of the Max Planck Institute for Psycholinguistics. They were paid for their participation. All of them were native speakers of Dutch and none of them reported any hearing deficit.

Procedure. Subjects were tested in groups of one to four in a quiet room. The stimuli were presented via headphones at a comfortable listening level. Subjects were asked to decide whether the last sound of each phrase was a [t] or a [k], and to press one of two response buttons, labelled “T” and “K”. Each subject heard each step of each sentence 12 times which resulted in a total number of 1152 sentences. The interval between items was 2500 ms. Four random lists including all experimental stimuli were constructed that varied with respect to the order of stimulus presentation. Those lists were distributed evenly across subjects so that each subject was presented with only one of the four lists. Because of the large number of experimental stimuli, subjects were tested in two sessions that were

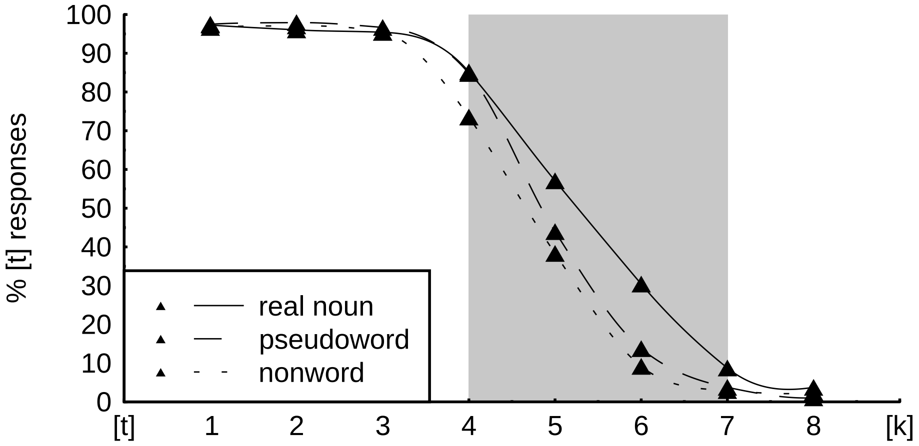


Figure 4.1: Experiment 4.1: Proportion [t] responses (in %) for each word type in the nominal sentence contexts on the [t]-[k] continuum (the ambiguous region is shaded in grey).

not further apart than one week. In each session they received 576 stimuli. In the second session, the response buttons were turned around so that reactions with each listener's preferred hand were evenly distributed over both [t]- and [k]-responses. The presentation of experimental items and the recording of the RTs was controlled by NESU software. The subjects were instructed to give their responses as fast as possible even if they felt uncertain about the sound. In each session subjects received a practice block consisting of 24 sentences before the real experiment started. Items were presented in three blocks within each session so that subjects were allowed two breaks. Each session lasted 50 minutes including breaks.

Results and Discussion

The percentage of [t] responses was computed for each subject as a function of sentence type and stimulus continuum. In Figure 4.1 the categorization functions of the three different word types presented in an NP are plotted across item sets (i.e., across mono- and bisyllabic final words). In Figure 4.2 the respective categorization functions for the VPs are plotted across item sets. Because a lexical shift typically shows up over a range of continuum steps rather than at a single cross-over point (see also Pitt & Samuel, 1993), a boundary region was chosen. The boundary region was defined as extending from step 4 to step 7. Analyses were carried out on the proportion of [t] responses in this boundary region. Following Miller and Dexter (1988), each listener's responses to each

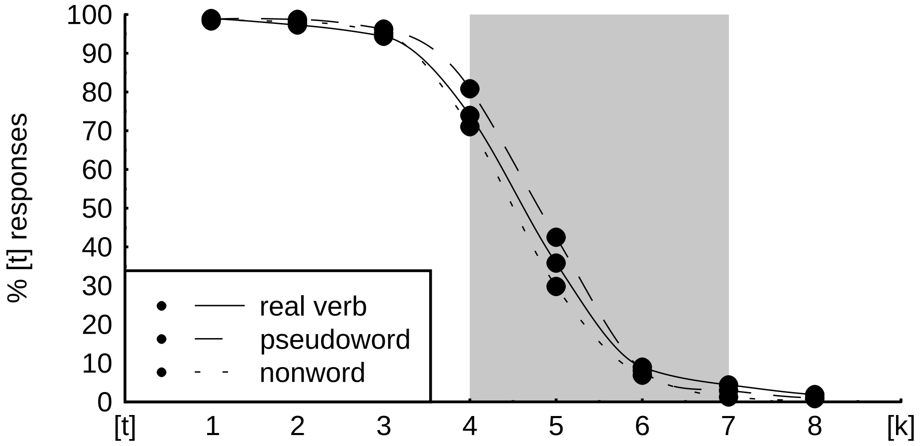


Figure 4.2: Experiment 4.1: Proportion [t] responses (in %) for each word type in the verbal sentence contexts on the [t]-[k] continuum.

step along the stimulus continuum were ranked and divided into three different reaction time groups: fast, medium, and slow RTs. Mean RTs (measured from the onset of the final sound) for these groups were respectively 350 ms (SD = 72 ms), 459 ms (SD = 85 ms), and 659 ms (SD 210 ms). The percentage of [t] responses to each stimulus in each RT range was then calculated for each subject as a function of context bias.

A one-way repeated-measures ANOVA on the proportion of [t] responses in the boundary region showed a significant effect of sentence type (whether the context predicted a noun or a verb; F(1,19) = 39.74, $p < .0001$) as well as a reliable lexical effect (whether the last word was a real word, a pseudoword, or a nonword; F(2,38) = 16.3, $p < .0001$). The interaction of these factors was also significant (F(2,38) = 22.18, $p < .0001$). The factor item set (whether the last word was mono- or bisyllabic) was significant (F(1,19) = 10.89, $p < .01$), as were the interactions of this factor with both the factor lexical status (F(2,38) = 4.99, $p < .01$) and the factor sentence type (F(2,38) = 48.22, $p < .0001$). There was also a three-way interaction of these factors (F(2,38) = 5.85, $p < .01$). The factor RT range (whether subjects' responses were ranked fast, medium, or slow) had a significant effect (F(2,38) = 7.6, $p < .01$). There was also an interaction of this factor with the factor lexicality (F(4,76) = 6.3, $p < .001$), the factor sentence-type (F(2,38) = 12.7, $p < .001$) and the factor item set (F(2,38) = 3.9, $p < .05$). None of the three-way interactions including the factor RT range was significant, nor was the four-way interaction significant. To summarize, the results suggest that the type of preceding sentence context played a role in categorization behaviour as did the lexical status of the final elements. The interaction of these factors indicates that preceding context only influenced categorization behaviour for a subset of final elements. Furthermore, phonemic decisions were determined by the set of items that were used but since this factor interacted with the other factors, this was again only true for a subset of the items. How fast subjects initiated their responses also had an influence on phonetic categorizations and the interactions of this factor with the other three factors indicate that lexicality, item set and sentence type influenced phonetic decisions differently over time.

For each RT range, one-way repeated measures ANOVAs were performed separately on the proportion of [t] responses in the boundary region. There was an effect of sentence type (whether the context predicted a noun or a verb) in the medium and slow RT ranges (medium: F(1,19) = 93.15, $p < .0001$; slow: F(1,19) = 6.89, $p < .05$) but no such effect in the fast RT range (F(1,19) = 0.33, n.s.). The lexical effect, however, was significant in all three RT ranges (fast: F(2,38) = 18.09, $p < .0001$; medium: F(2,38) = 8.29, $p < .01$; slow: F(2,38) = 8.29, $p < .01$). There was a significant interaction of these factors in the fast and medium RT ranges (fast: F(2,38) = 4.62, $p < .05$; medium: F(2,38) = 9.07, $p < .001$). The factor item set was not significant in the fast RT range but was significant in both the medium (F(1,19) = 22.24, $p < .0001$) and the slow (F(1,19) = 10.15, $p < .01$) RT ranges. In the fast RT range the interaction of that factor with the factor context type was significant (F(1,19) = 16.85, $p < .001$) as was the three-way interaction of these two factors with the factor lexicality (F(2,38) = 6.06, $p < .01$). None of these interactions were significant in the medium or the slow RT ranges.

The interaction of the factors sentence type and lexical status observed in the overall categorization data was due to listeners' fast and medium responses. The three-way interaction of the former two factors with the factor item set was confined to listeners' fastest responses while the factors sentence type and item set influenced only the medium and slower responses.

Noun Phrases

In order to examine this complex pattern in more detail, individual t-tests compared the lexical-status conditions within each sentence context (NP vs. VP). These pairwise comparisons were also performed within each RT range. Figure 4.3 shows the proportions of [t] responses for each word-type in the ambiguous region of the continuum in each RT range for the nominal sentence contexts while Figure 4.4 illustrates the respective proportions separately for each item set (i.e., mono- vs. bisyllabic final words). T-tests showed that, overall, all pairwise comparisons between the three lexical-status conditions within the NP contexts differed significantly from each other. Sentences that ended with a real noun (e.g., ’straat’) got significantly more [t] responses in the ambiguous area than both pseudowords (e.g., ’vlaat’; t(19) = 4.21, p < .0001) and nonwords (e.g., ’klaat’; t(19) = 5.93, p < .0001). Pseudowords in this context got significantly more [t] responses as compared to nonwords (t(19) = 3.19, p < .01).

The lexical effect for nouns was then examined in each RT range. Nouns in the fast RT range got significantly more [t] responses as compared to the other two conditions (nouns vs. pseudowords: t(19) = 4.41, p < .0001; nouns vs. nonwords: t(19) = 4.66, p < .0001) while there was no difference between the latter two conditions (t(19) = 0.77, n.s.). Whereas pseudowords in the fast RT range did not produce more [t] responses than nonwords, they started to diverge from nonwords in the medium and slow RT ranges (medium: t(19) = 3.21, p < .01; slow: t(19) = 3.41, p < .01) and were no longer different from real nouns (medium: t(19) = 0.80, n.s.; slow: t(19) = 0.92, n.s.). The function for real nouns, however, showed a robust lexical shift throughout the whole time range as compared to nonwords. Thus, in the medium and slow RT ranges, listeners also labelled an ambiguous sound more often as [t] when this sound formed a real noun than when it formed a nonword (medium: t(19) = 3.43, p < .01; slow: t(19) = 3.16, p < .01).

Because of the significant interactions of the factor item set (mono- vs. bisyllabic final words) with both sentence type and lexical status, the data were split by item sets and the three lexical-status conditions within each set were compared in separate t-tests. This procedure revealed different patterns for the mono- and the bisyllabic item sets (see Figure 4.4). In the monosyllabic item set the categorization function for the pseudoword ’vlaat’ clearly clustered with that for the nonword ’klaat’. Both conditions produced significantly fewer [t] responses as compared to the real noun ’straat’ but did not differ from each other (*straat* vs. *vlaat*: t(19) = 6.58, p < .0001; *straat* vs. *klaat*: t(19) = 5.06, p < .001; *vlaat* vs. *klaat*: t(19) = 0.25, n.s.). In the bisyllabic item set, however, the pseudoword ’magiet’ seems to have been treated exactly like the real noun ’bandiet’ as both conditions produced more [t] responses than the nonword ’keliet’ (*bandiet* vs. *magiet*: t(19) = 0.14, n.s.; *bandiet* vs. *keliet*: t(19) = 3.5, p < .01; *magiet* vs. *keliet*: t(19) = 3.85, p < .001).

The factor item set was then examined in each RT range (see Figure 4.4). For both item sets the patterns in the medium RT range very much resembled those in the fast RT range: the pseudoword ’vlaat’ in the monosyllabic item set was treated like the nonword ’klaat’ (both received significantly fewer [t] responses than the noun ’straat’) while the pseudoword ’magiet’ in the bisyllabic item set was treated like the noun ’bandiet’ (both received more [t] responses than the nonword ’keliet’ although only the pseudoword-nonword comparison was fully

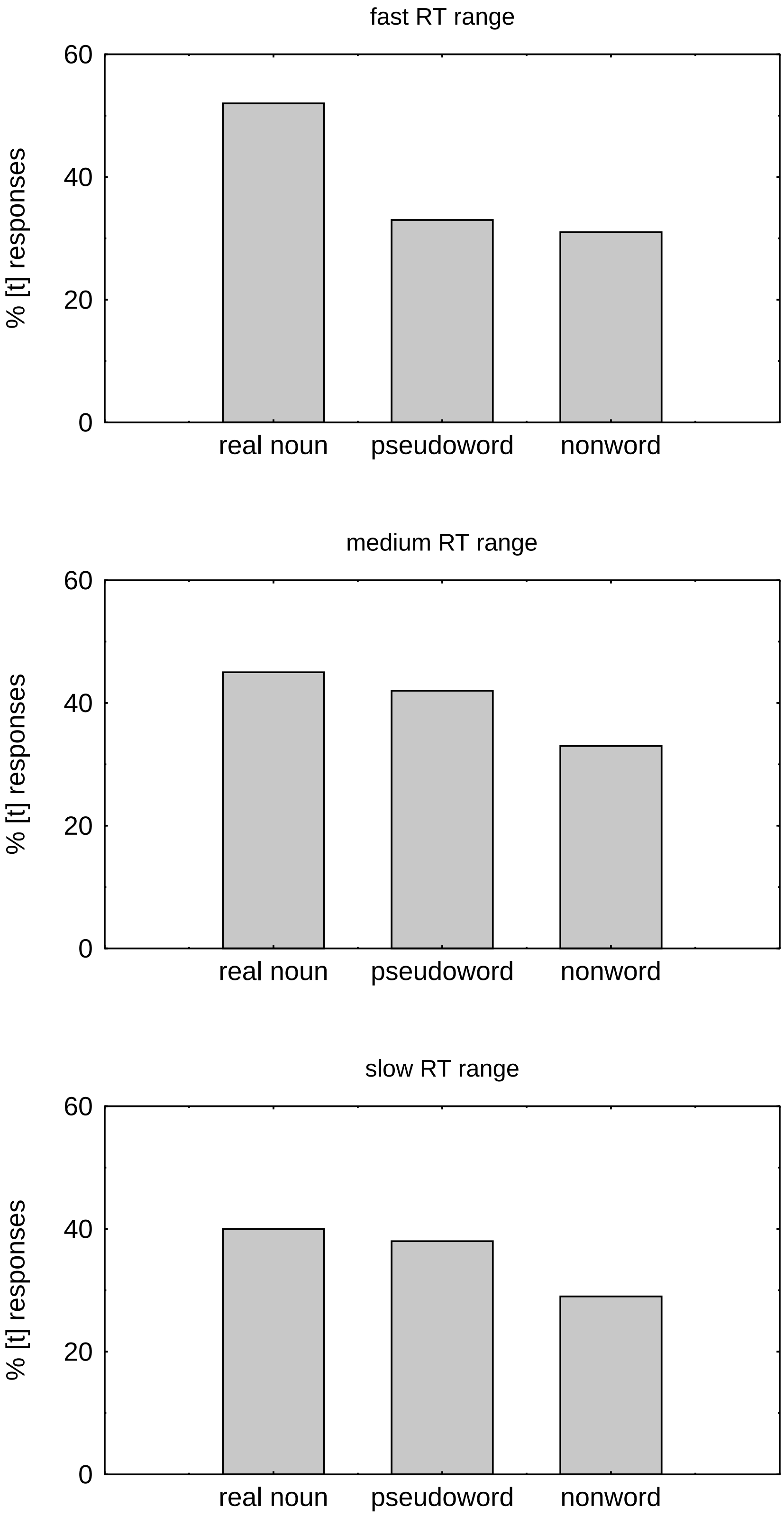
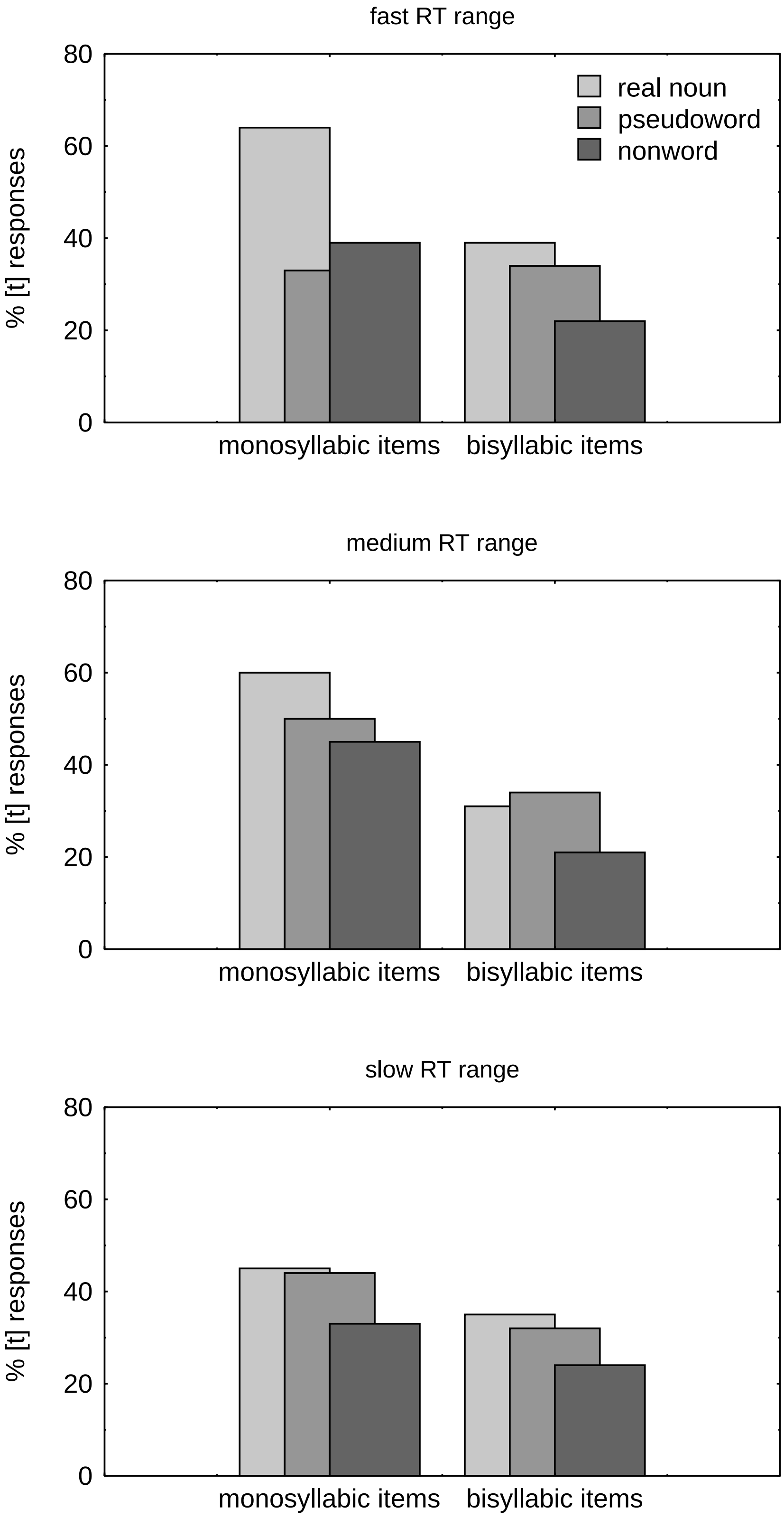


Figure 4.3: Experiment 4.1: Proportion [t] responses (in %) for each word type in the nominal sentence contexts for each RT range.


fast RT range

medium RT range

slow RT range

Figure 4.4: Experiment 4.1: Proportion [t] responses (in %) for each word type in the nominal sentence contexts for each RT range plotted separately for each item set.

reliable). For the bisyllabic item set exactly the same pattern was also observed in the slow RT range (although this time only the noun-nonword comparison was fully reliable). The pattern for the monosyllabic item set, however, changed slightly in listeners' slowest responses: while the noun 'straat' was still significantly different from the nonword 'klaat' the noun-pseudoword (i.e., 'straat-vlaat') comparison was no longer significant while the pseudoword 'vlaat' received significantly more [t] responses than the nonword 'klaat'.

A possible explanation for the different patterns between the two item sets might be the following: the Adjectival Phrase (AP) 'een stoere ...' (*a cool* ...) that served as preceding context in the bisyllabic item set to form an NP with the final word 'bandiet' (*a cool bandit*) could also be interpreted as an NP (*a macho*) when read in isolation or when followed by a verb like, for example, 'rookt' (*smokes*) yielding the VP 'een stoere rookt' (*a macho smokes*). Although the string 'magiet' is not a Dutch word it can be decomposed into the morphemes 'magie' and '-t' (as already outlined in the Introduction) which would in combination with the NP 'een stoere' form the VP *a macho magics*. This sort of NP-interpretation, however, is not possible (or much less likely) for the AP 'een brede' (*a wide*) which was used in the monosyllabic item set. Thus, a phrase like '*een brede rookt' (*a wide smokes*) is ungrammatical. Therefore it is unlikely that listeners interpreted the phrase 'een brede vlaat' as a VP (*a wide custards*) rather than as an NP. That might explain why the pseudoword 'magiet' was treated more like a real word while the pseudoword 'vlaat' was treated more like a nonword. This explanation of course presupposes that listeners not only were aware of the morphological components of the final string 'magiet' but also used that information for judging the ambiguous phonemes. This hypothesis is supported by a similar effect for the pseudoword 'magiet' presented in a verbal context 'de moeder magiet' which will be discussed in more detail below. Overall, there were more [t] responses to the pseudoword 'magiet' *magics* as compared to both the real verb 'geniet' *enjoys* (t(19) = -1.99, $p < .06$) and the nonword 'keliet' (t(19) = 3.4, $p < .01$), although only the latter difference was fully reliable.

Verb Phrases

The pattern for VP contexts was different from the one observed for noun phrases. Figures 4.5 and 4.6 show the proportions of [t] responses for each word type in the ambiguous region of the continuum in each RT range for the verbal sentence contexts across item sets (4.5) and separately for each item set (4.6) respectively. Overall, *pseudowords* (*vlaat* and *magiet*) received the most [t] responses in the ambiguous region (see also Figure 4.2). The comparison with the nonwords (*klaat* and *keliet*) was reliable (t(19) = 4.01, $p < .01$) while the one with real verbs (*gaat* and *geniet*) was not (t(19) = 1.75, n.s.). Real verbs received significantly more [t] responses as compared to nonwords (t(19) = 2.17, $p < .05$). As for the noun phrases, the effects were further examined in each RT range. In the fast RT range the real verbs and the pseudowords received more [t] responses as compared to the nonwords (real verbs vs. nonwords: t(19) = 2.8, $p < .01$; pseudowords vs. nonwords: t(19) = 2.44, $p < .05$) while there was no difference between the former two conditions (t(19) = 1.66, n.s.). This pattern changes, however, in the medium RT range where the pseudowords received the most [t] responses as compared to both real verbs (t(19) = -4.7, $p < .0001$) and nonwords (t(19) = 3.48, $p < .01$). As can also be seen in Figure 4.5 the lexicality effect for real verbs that was there in the fast RT range was absent in the medium RT range (real verbs vs. nonwords: t(19) = -0.98, n.s.). Although in Figure 4.5 it looks as if this lexicality effect emerges again in the slow RT range, none of the pairwise comparisons yielded significant effects (real verbs vs. pseudowords: t(19) = -0.98, n.s.; real verbs vs. nonwords: t(19) = 1.39, n.s.; pseudowords vs. nonwords: t(19) = 1.86, n.s.).

As in the NP-analyses, different patterns were observed for the two item sets (monosyllabic vs. bisyllabic final words). In the monosyllabic item set with the preceding context ’de tante’ (*the aunt*), overall there was a weak bias towards the [t] endpoint for the pseudoword ’vlaat’ *custards* as compared to the nonword ’klaat’ (t(19) = 2.03, $p < .06$) while no other pairwise comparison showed reliable effects (*vlaat* vs. *gaat*: t(19) = -0.67, n.s.; *gaat* vs. *klaat*: t(19) = 1.22, n.s.). Although the real verb ’gaat’ received more [t] responses than both the pseudoword ’vlaat’ and the nonword ’klaat’ in the fast RT range, this tendency was not significant. The weak bias towards the [t] endpoint for the pseudoword ’vlaat’ observed in the overall analysis was due to a marginal effect in the medium RT range (*vlaat* vs. *klaat*: t(19) = 1.95, $p < .07$). Also the pseudoword-word comparison was only marginally significant in the medium RT range (*vlaat* vs. *gaat*: t(19) = 1.85, $p < .08$). In the slow RT range, the pseudoword ’vlaat’ still received the most [t] responses as compared to the other two conditions but the comparison with the real verb ’gaat’ was not even marginally significant while the comparison with the nonword klaat was (*vlaat* vs. *gaat*: t(19) = 1.16, n.s.; *vlaat* vs. *klaat*: t(19) = 2.04, $p < .06$). Also the pairwise comparison of the real word ’gaat’ and the nonword ’klaat’ was marginally significant whereby the latter received less [t] responses (*gaat* vs. *klaat*: t(19) = 1.82, $p < .09$).

As already mentioned above (in the section on noun phrases), there were more [t] responses in the bisyllabic item set to the pseudoword ’magiet’ (*magics*) as compared to both the real verb ’geniet’ (*enjoys*) (t(19) = -1.99, $p < .06$) and the nonword ’keliet’ (t(19) = 3.4, $p < .01$) although only the latter difference was fully

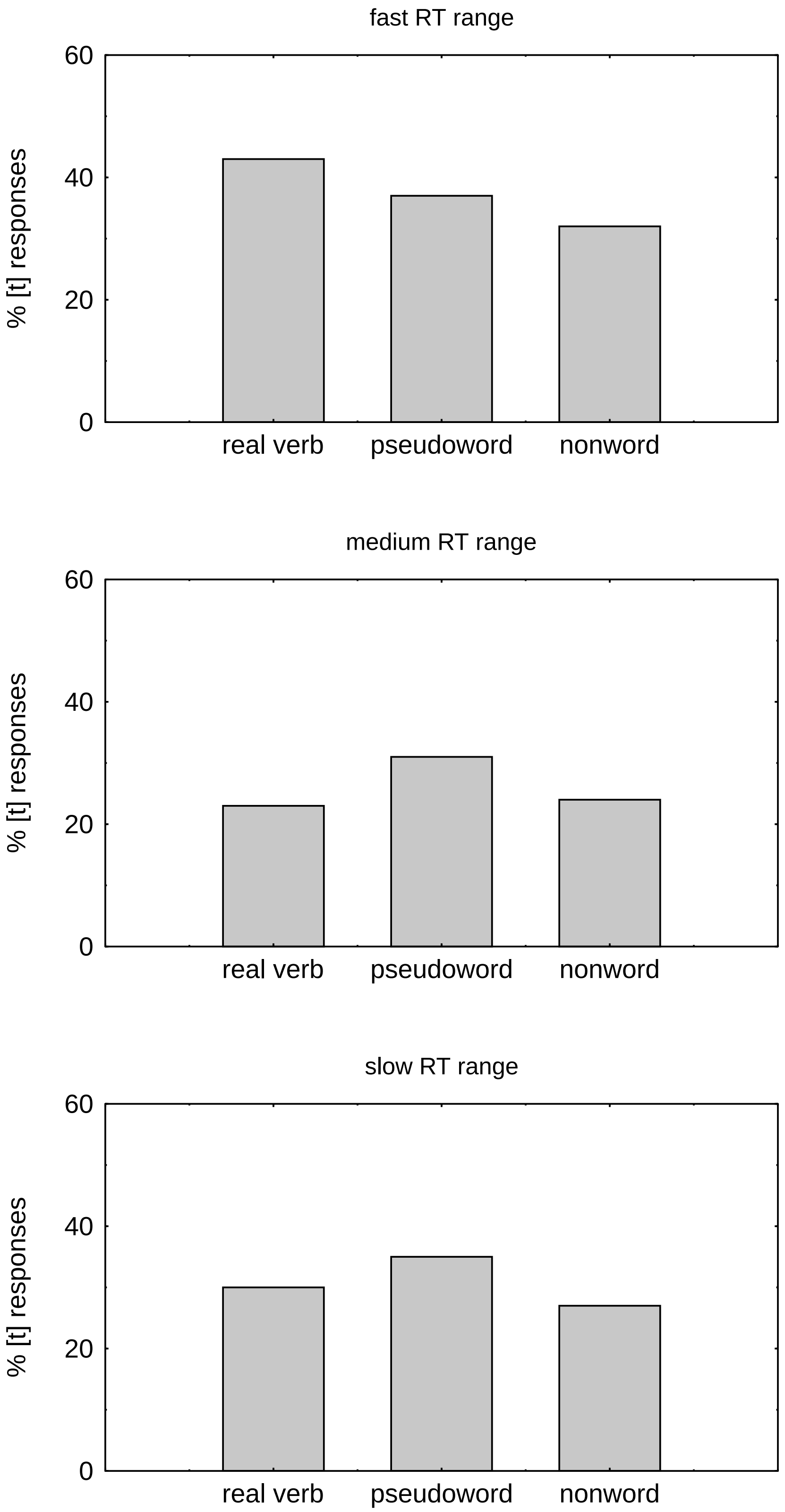


Figure 4.5: Experiment 4.1: Proportion [t] responses (in %) for each word type in the verbal sentence contexts for each RT range.

reliable. This ’lexical’ effect for the pseudoword ’magiet’ (i.e., the pseudoword-nonword comparison) stayed stable over the fast and medium RT ranges, while a lexical effect for the real verb ’geniet’ only emerged in the fast RT range (*geniet* vs. *keliet*: t(19) = 2.73, $p < .05$). As already argued above, the explanation for the overall ’lexical’ effect for the pseudoword ’magiet’ automatically implies that listeners could interpret the complex string ’magiet’ as a morphologically complex (though nonexistent) form and integrate it into the context ’de moeder magiet’. Otherwise there should be no shift in the categorization function for that phrase.

Two independent confounds might be responsible for the missing or weak lexical effects for real verbs in both item sets. In the bisyllabic item set the VP ’de moeder geniet’ (**the mother enjoys*) - which is ungrammatical in English - is somewhat odd in Dutch. Although strictly speaking the phrase is syntactically correct, there is a strong expectation for a following Object-NP to define what the mother enjoys. This might have prevented the sentential and lexical information from exerting a strong influence on listeners’ decisions. Note, however, that there was a lexical effect in the fast RT range for that verb which then vanished in the two slower RT ranges so that the overall effect did not reach significance.

Another explanation holds for the lack of a lexical effect (or in fact the lack of a bias towards [t] in any context) in the monosyllabic item set. In this item set listeners always heard two instances of a clear [t] in the preceding context (i.e., de **t**an**t**e gaat) just before they had to label the final sound. The lack of a lexical and/or sentential effects might be due to a contrast effect. Because listeners had just heard the unambiguous [t] sounds in the preceding word ’tante’, they may have been less likely to accept a following ambiguous sound as a [t] (see for more information on contrast effects Alfonso, 1981 and Fox, 1984).

Conclusions

Two major findings can be reported. First, the pattern for verb contexts was different from the pattern for noun contexts. While overall there was a clear bias towards the [t] endpoint for real nouns in final positions of NPs, this bias was less obvious for verbs that appear at the ends of VPs. Also, the lexical effect for nouns was more stable across the time course of responding. While the effect for both nouns and verbs became smaller in slower responses, the effect for nouns was significant in all three RT ranges, while the effect for verbs was only reliable in fast responses and disappeared from the medium time range on. Second, the pseudowords showed different categorization functions as a reflection of the preceding context. When presented in the final position of VPs, pseudowords showed the strongest bias towards the [t] endpoint, as compared to real verbs and nonwords presented in the same sentential context. They differed from non-

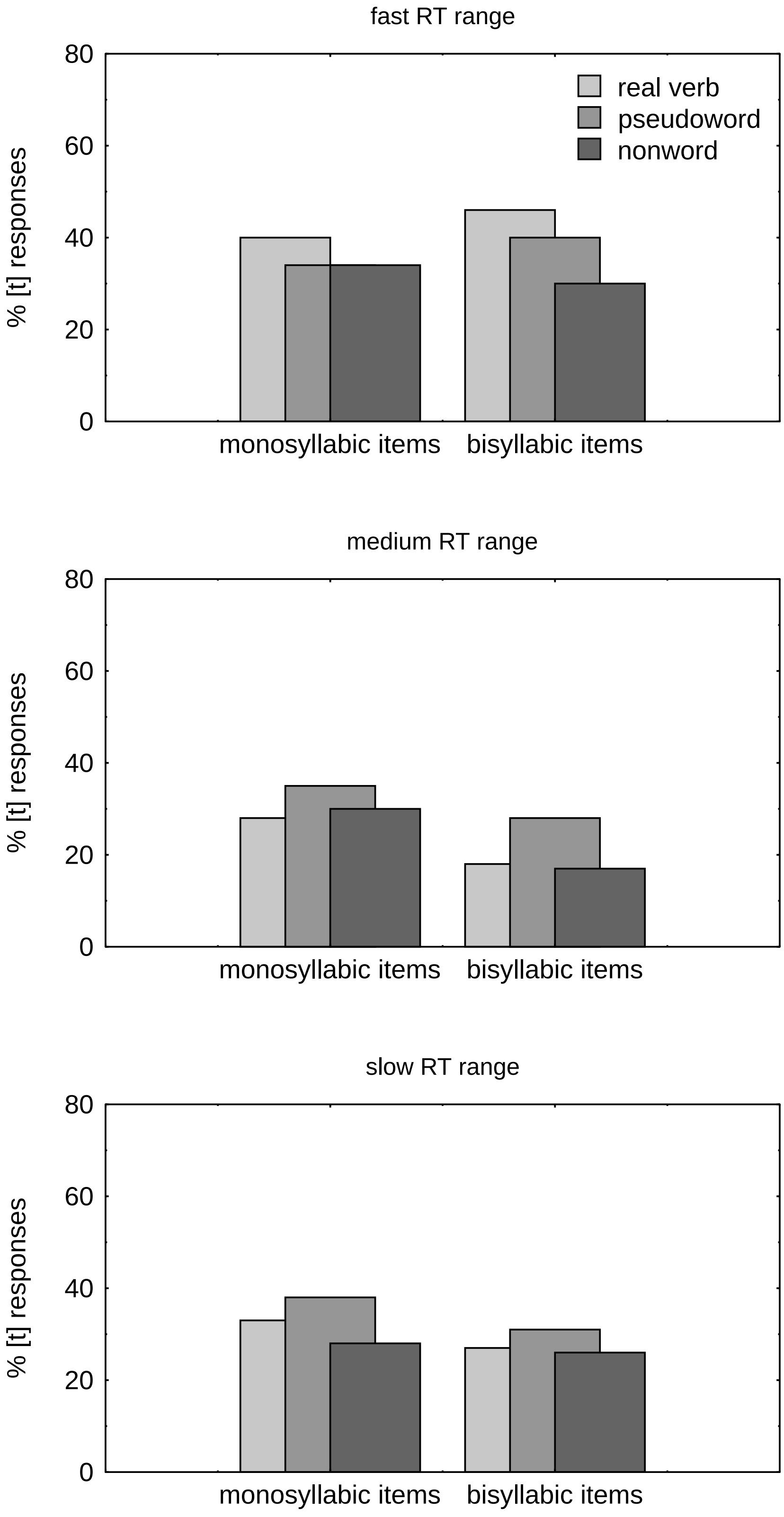


Figure 4.6: Experiment 4.1: Proportion [t] responses (in %) for each word type in the verbal sentence contexts for each RT range plotted separately for each item set.

words up until the medium RT range while they differed from the real verbs only in the medium RT range. When pseudowords were presented at the final position of NPs, however, their categorization function lay right between those for real nouns and nonwords. Over time, the ’lexical effect’ for pseudowords built up: it was absent in the fast RT range but was present in both the medium and the slow RT ranges.

As outlined above, some confounds in the stimulus design might have been responsible for the current results. The data are therefore too inconsistent to allow one to draw any strong conclusions about the questions of interest. Nevertheless, the results were encouraging enough to warrant further investigations. The categorization functions for the pseudoword ’magiet’ in both NP- and VP-contexts, for example, might be interpreted as a hint of a morphological effect in phonetic categorization. In both contexts, the strongest bias towards the [t] endpoint was observed for the pseudoword. This similarity suggested that the NP ’een stoere magiet’ might have been interpreted as a VP rather than an NP. The pattern of categorization functions for monosyllabic final items in the NP contexts (i.e., *een brede straat, vlaat, klaat*) support this interpretation: there was no bias towards more [t] responses for the pseudoword ’vlaat’ since ’een brede’ is not a possible NP.

If this interpretation were true, it should be possible to establish a *morphological* effect for real verbs as well. Due to the confounds outlined above, this sentential or even morphological effect could not be observed with the current stimuli. New phonetic categorization experiments with more carefully designed stimuli were therefore designed. They are presented in Chapter 5.

The role of morphology in phonetic decision-making - Part II

Chapter 5

Due to some confounds in the materials used in Experiment 4.1, the question whether in phonetic categorization ambiguous sounds as parts of inflected verbs might be treated differently than the same ambiguous sounds as parts of noun stems could not be answered satisfactorily. While inflectional morphemes (like the ’-t’ in *gaat*) are syntactically predictable from the preceding sentential context, phonemes that are part of noun stems (like the ’-t’ in *straat*) are not. Therefore it was hypothesised that different biases in the categorization functions for inflected verbs as compared to uninflected nouns might reflect that difference. What might these different biases look like? One possibility is that a sound ambiguous between [t] and [k] at the end of the final verb ’gaa?’ in a verbal phrase like ’de tante gaa?’ is more likely to be labelled as [t] than the same ambiguous sound at the end of a nominal phrase like ’een brede straa?’. In the former case *three* processing levels - the phonemic, the lexical, and the syntactic level - might bias the phonetic decision, while only two of these levels - the phonemic and the lexical levels - should influence the decision in the latter case. The result might be a stronger bias towards [t] when the ambiguous sound is a potential morpheme than when this sound is part of a noun stem.

Another possibility might be that there were more [t] responses to ambiguous sounds in uninflected nouns than to the same ambiguous sounds in inflected verbs. Because the inflectional marker ’-t’ can be more easily separated from the rest of the sentence than a noninflectional ’-t’ (i.e., ’-t’ in *straat*), listeners might be able to *ignore* preceding context and categorize sounds in morphemic position independently of the context.

It is of course also possible that noun phrases and verbal phrases will show similar categorization patterns overall but reveal different time patterns of the effects. It might be that syntactic information is processed earlier than semantic information and therefore exerts its influence only on fast phonetic decisions while semantic information might be effective somewhat later. As in Experiment

4.1, the time course of categorization will again be analysed.

In order to deal with the confounds reported in Chapter 4, new sets of materials were created. The experimental sentences were chosen such that the potential (mis)interpretation of APs like *een brede* in Experiment 4.1 as NPs could be avoided. Therefore, full sentences were used rather than minimal phrases. Because Dutch is a *verb second* language, questions were constructed so that sentence-final presentation of verbs was legal. Because Experiment 4.1 showed that ambiguous sounds in pseudowords like ’magie?’ were categorized more often as [t] than the same sounds in nonwords like ’kelie?’ it is possible that listeners decomposed the material in some way and used that information to label the sounds. Therefore, pseudowords were also included in the new materials. Experiment 5.1A will test real verbs and nouns and compare those to matched nonwords in the same sentential contexts while pseudowords will be tested against the same nonwords in Experiment 5.1B.

Experiment 5.1A: Verbs and nouns vs. nonwords

Method

Materials: Experiments 5.1 - 5.3. The questions used in Experiments 5.1A and 5.1B either ended with a verb (A) or a noun (B).

(A: VP) Vraag jij of Jan morgen gaat?
Are you asking whether Jan leaves tomorrow?

(B: NP) Zie jij nog wel eens een plaat?
Do you see a record now and then?

Two sets of final items were chosen: one consisted of monosyllabic and the other of bisyllabic words. The mean surface frequencies of the verbs and the nouns were 7462 (per 42 million words based on the CELEX computerised database of Dutch) and 1109 (per 42 million words) respectively. The preceding contexts were designed such that no word included either a [t] or a [k] in order to avoid any contrast effects.

As in Experiment 4.1, each [t]-final item within each set was paired with a [k]-final nonword so that four different continua were formed: verb-nonword (e.g., *gaat - gaak*), noun-nonword (e.g., *plaat - plaak*), pseudoword-nonword (e.g., *vlaat - vlaak*), and nonword-nonword (e.g., *snaat - snaak*). The four different types of final words (verb, noun, pseudoword, and nonword) were matched phonologically such that each of them contained the same vowel before the final sound. Each nonword-nonword and pseudoword-nonword continuum was presented in both the verbal and the nominal contexts while each verb-nonword and noun-nonword continuum was only presented in the grammatically appropriate sentence contexts. To increase the variation of items, and by that increase the subjects' attention to the sentence contexts, 8 filler sentences were constructed. Half of these ended with a verb-nonword continuum and half ended with a noun-nonword continuum. Each filler-sentence was also presented with a phonologically-matched sentence-final nonword-nonword continuum. For a full list of all sentences, see Appendix D.

Stimulus construction. All sentences were recorded by a female native speaker of Dutch (who had already spoken the items for Experiments 2.1 - 3.3, and 4.1). The technical procedure was exactly the same as in Experiment 4.1. Again, a fifteen-step place-of-articulation continuum was constructed. For this, a [t] and a [k] were spliced out of natural waveforms from the verb 'gaat' (*leaves*) and the matched nonword 'gaak', with the cuts being made at zero crossings. Both stimuli were spliced such that they were 212 ms long. The thirteen intermediate stimuli were constructed using the same procedure as in Experiment 4.1 (Stevenson, 1979; Repp, 1981). Again the vowel preceding the last sound was kept constant within each item set. In order to keep the transitions in the preceding vowel constant so that any bias towards [t] or [k] would be the same for each word or nonword, the same vowels were used in all contexts. As in Experiment 4.1, vowels were taken from velar contexts and had therefore transitions consistent with a [k]. Any bias based on these transitions would therefore work against the predictions. The fifteen tokens from the [t]-[k]-continuum were then spliced onto all words. These words were than spliced into the sentence contexts. Both the nonword-nonword continua and the pseudoword-nonword continua were the same in the verbal and the nominal contexts.

As in Experiment 4.1, a pretest was conducted to establish the ambiguous region of the fifteen-step continuum. Eight native Dutch listeners were presented with a subset of the whole range of experimental stimuli and were asked to label the sentence-final sounds as either [t] or [k]. While all verb-nonword and noun-nonword continua were presented in the respective sentence contexts, the pseudoword-nonword and nonword-nonword continua appeared only in one of the two contexts. Thus, eight different sentences were presented in the pretest. Each of the fifteen steps was presented only once in each of the eight sentences so that a total of 120 stimuli was presented to each listener. Visual inspection of the categorization functions showed that steps 1 - 5 and 12 - 15 were clear instantiations of a [t] and a [k] respectively, while steps 6 - 11 were ambiguous. Steps 5 - 12 were thus chosen for presentation in the experiment and will henceforth be called steps 1 - 8.

Subjects. Twenty student volunteers (15 female and 5 male) from the Max Planck Institute subject pool were paid for their participation. None of them reported any hearing impairment.

Procedure. Four different randomizations were constructed that all contained the same number of experimental stimuli and filler sentences. They only differed with respect to the order of presentation. The lists were evenly distributed across subjects. Each subject heard each step of each sentence 12 times which resulted in a total number of 768 experimental sentences. Each step of the continuum for each of the filler sentences was only presented once which summed up to a total number of 128 filler presentations. The interval between items was 2500 ms. Subjects were tested in groups of one to three in sound-attenuated booths. The technical setup and the facilities were the same as in Experiment 4.1. Subjects had to come back for a second session due to the high number of stimulus presentations. Between the sessions there was a maximum period of one week. In each session listeners heard 24 practice sentences that were then followed by three experimental blocks that contained 448 trials in total. Instructions were the same as in Experiment 4.1. As in that experiment, response buttons were turned around in the second session so that reactions with each listener's preferred hand were balanced across the two phonemes. Sessions did not last longer than 50 minutes including breaks.

Results and Discussion

For each subject the percentage of [t] responses was computed as a function of sentence type and stimulus continuum. Figure 5.1 shows the categorization functions of the four different sentence types across item sets. A boundary region was chosen which extended from step 4 to step 7. All analyses were carried out on the proportion of [t] responses in that area. Each listener's responses to each step along the stimulus continuum were then ranked and divided into fast, medium, and slow reaction time groups. The mean RTs (measured from the onset of the final sound) for these groups were 417 ms (SD = 126 ms) in the fast, 545 ms (SD = 156 ms) in the medium and 779 ms (SD = 299 ms) in the slow RT ranges. For each subject the percentage of [t] responses to each stimulus in each RT range was calculated as a function of context bias.

The result of a one-way repeated-measures ANOVA on the proportion of [t] re-

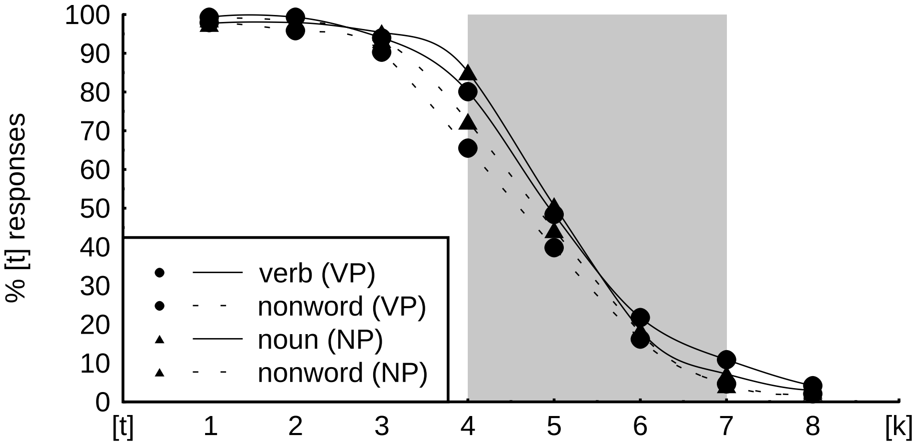


Figure 5.1: Experiment 5.1A: Proportion [t] responses (in %) for each sentence type on the [t]-[k] continuum.

sponses in the boundary region revealed a lexical effect (whether the last word was a real word or not; F(1,19) = 30.0, $p < .0001$) but no effect of the factor context type (whether the context predicted a noun or a verb) and no interaction of these factors. There was a main effect of the factor item set (whether the final word was mono- or bisyllabic; F(1,19) = 31.3, $p < .0001$) but no interaction of this factor with either lexicality or sentence type. The factor RT range (whether responses were ranked fast, medium, or slow) was also significant (F(2,38) = 3.2, $p < .05$) as were the interactions of this factor with the factor lexicality (F(2,38) = 3.6, $p < .05$) and the factor item set (F(2,38) = 27.8, $p < .0001$). Also the three-way interaction of the factors RT range, lexicality and item set was significant (F(2,38) = 3.2, $p < .05$) while the three-way interaction of RT range with lexicality and context type was only marginally significant (F(2,38) = 3.2, $p < .06$). None of the other interactions reached significance.

Post-hoc t-tests showed that both the real verbs and the real nouns received significantly more [t] responses than their nonword counterparts (verb vs. nonword in VPs: t(19) = 4.59, $p < .001$; noun vs. nonword in NPs; t(19) = 4.43, $p < .001$). The significant influence of the factor item set (mono- vs. bisyllabic final words) was exclusively due to an overall stronger bias towards the [k] endpoint (i.e., less [t] responses) for the monosyllabic item set as compared to the bisyllabic item set. This stronger bias towards [k] can be explained in the following way: all final words or nonwords in the monosyllabic item set contained the vowel [ɑ] while all words and nonwords in the bisyllabic item set contained the vowel [i]. Both vowels were spliced from velar contexts and had thus transitions towards a [k]. Earlier research (e.g., Smits, 2000) has demonstrated that formant transitions in the vowel [ɑ] can be perceived reliably better than the same formant transitions in the vowel [i]. The transitions towards a [k] were thus more detectable in the monosyllabic item set than in the bisyllabic item set which resulted in an overall bias towards the [k] endpoint for the monosyllabic as compared to the bisyllabic item set.

As suggested by the missing interactions with the factors lexicality and sentence type, other than this overall difference between the two item sets there were no different patterns between them. The interaction of this factor with RT range was significant because the overall [t] bias for bisyllabic items was significant in the fast and medium RT ranges (*fast*: F(1,19) = 46, $p < .0001$; *medium*: F(1,19) = 13.7, $p < .01$) but vanished in the slow RT range. In further analyses the factor item set will not be taken into account.

Separate one-way repeated measures ANOVAs were performed on the proportion of [t] responses in the boundary region of each RT range. The respective proportions are illustrated in Figure 5.2. In the fast RT range there was a main effect of lexicality (F(1,19) = 27.82, $p < .001$) and a significant interaction of lexicality with the factor context type (F(1,10) = 8.15, $p < .01$) while the main effect of context type was not significant. A t-test showed a significant difference between real verbs and their nonword counterparts (t(19) = 5.98, $p < .001$), while there was no such effect for the nominal item set. The lexical effect was still significant in the medium RT range (F(1,19) = 11.27, $p < .01$), but the interaction between lexicality and context type vanished. As revealed by a post-hoc t-test the lexical effect here was due to a significant difference between real nouns and their nonword counterparts (t(19) = 2.37, $p < .05$). There was no significant difference between real verbs and the relevant nonword condition. In the slow RT range the lexical effect disappeared and also pairwise post-hoc comparisons indicated that there were no significant effects.

In sum, there are clear indications that the functions for real verbs and nouns as compared to the respective nonword conditions showed different patterns over time. The lexical effect for verbs was strongest in the fast RT range and became weaker over time while the lexical effect for nouns was only significant in the medium RT range.

The different patterns for the verbal and nominal conditions suggest that listeners were sensitive to morphological information during phonological decision-making. Especially when responding fast, they seemed to profit from the fact that the phoneme they had to judge had a morphological status. Furthermore, the time course of the categorization bias also implies that the influence of the sentential information on the categorization of a potential morpheme is tempo-

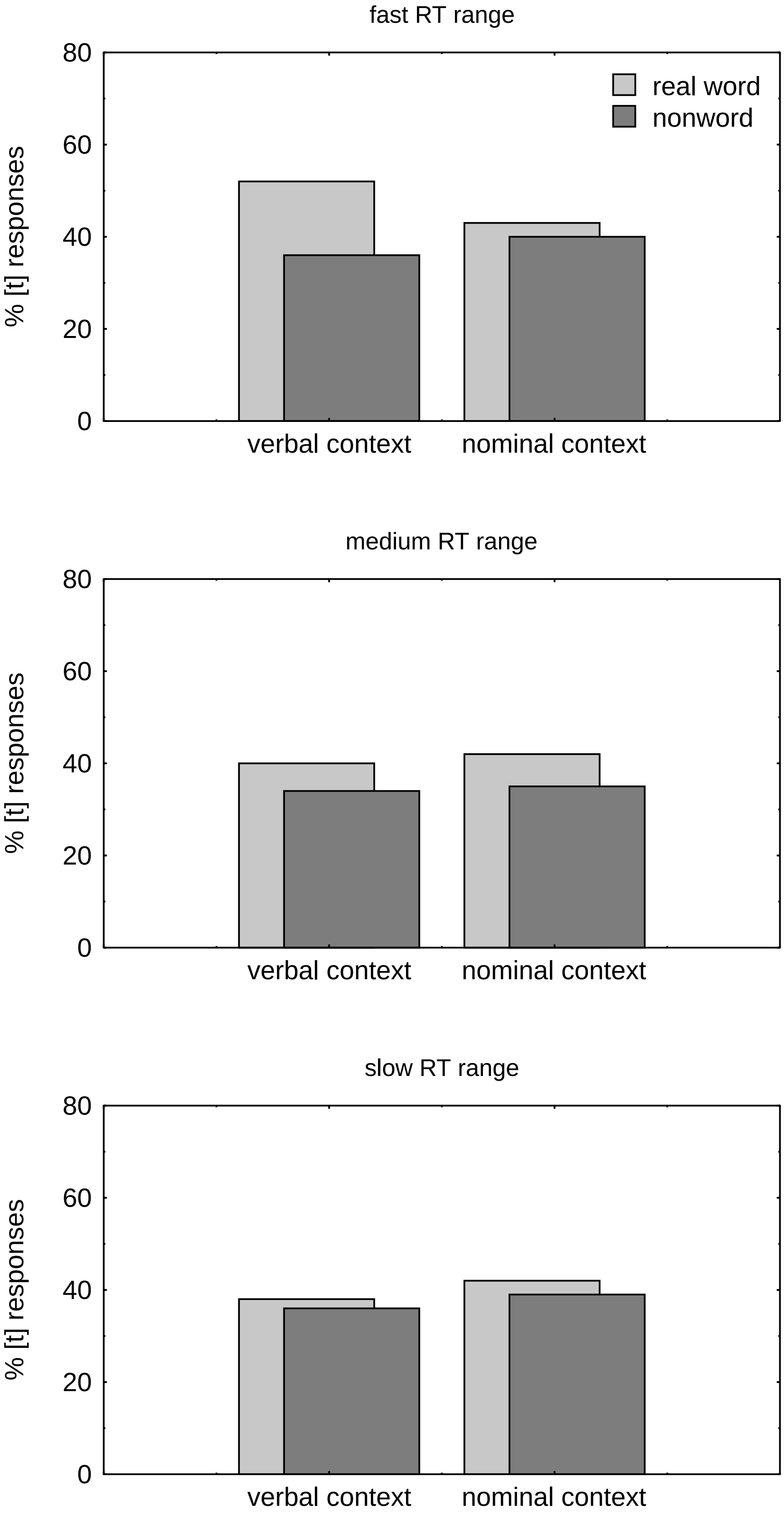


Figure 5.2: Experiment 5.1A: Proportion of [t] responses (in %) in the ambiguous continuum region for each sentence type in each RT range.

rally restricted. Only when responding fast did people benefit from this source of information. This is in line the predictions and with previous findings on syntactic sentential influences on phonetic categorization (see van Alphen & McQueen, 2001). Once the integration of morpho-syntactic information has been completed, this information can no longer influence phonemic decisions. Therefore only the fastest responses were affected by that sort of information.

But why did the lexical effect for nouns not show a similar pattern? In this case, only the responses in the medium RT range were affected by sentential information. One possible explanation might be that during sentence processing integration of semantic information happens somewhat later than syntactic integration. If so, this information would be accessible later than syntactic information and would therefore influence phonetic decisions in slower responses rather than in the fastest responses. But, just like syntactic information, once this information has been fully integrated it is not accessible anymore to influence the slowest responses. This interpretation is furthermore in line with the results of the verbal condition where also semantic interpretation is required. In the medium RT range there is a tendency of listeners to label ambiguous sounds more as [t] than [k] when they were parts of real verbs. This might be a weak semantic effect since semantic integration is of course also required in the verbal phrases.

There is, however, one other interpretation for the stronger lexical effect that was observed for inflected verbs. Remember that the mean surface frequency of the verbs (7264) was much higher than the mean surface frequency of the nouns (1109). It is therefore possible that the different effects observed for inflected verbs and uninflected nouns is a frequency effect instead of a reflection of the processing of the preceding context. Frequency can influence listeners' phonetic decision-making (Connine, Mullennix, Shernoff, & Yelen, 1990; Connine, Titone, & Wang, 1993). These authors reported more responses in the ambiguous region of voicing continua to the continuum endpoint which formed the more frequent word. But if the current pattern of results is a pure frequency effect the same pattern should be obtained when the same words are presented in isolation. That is, inflected verbs should receive overall more [t] responses in the ambiguous region than nouns. This bias should be strongest in the fast RT range while the lexical effect for nouns should be strongest in the medium RT range.

Thus, in order to be able to attribute the current results to the processing of the sentence contexts it is necessary to demonstrate that the different effects found for nominal and verbal phrases vanish when the same final words are presented in isolation. If the differential effects, however, are observed for inflected verbs and uninflected nouns in isolation, this would mean that there are more general access differences between the two categories (probably due to frequency differences).

The final items were presented in isolation in Experiment 5.2. The second part of Experiment 5.1 will be reported first. This experiment examined the case of pseudowords presented in different sentential contexts. Remember that pseudowords form a special case of the role morphology might play in phonetic decision-making: would there be more [t] responses to ambiguous sounds when, on the basis of the preceding context, the decomposition of a pseudoword like ’vlaat’ into the noun ’vla’ and the inflectional marker ’-t’ would make sense (as in ’Vraag je of Jan morgen vlaat?’, *Are you asking whether Jan custards tomorrow?*) than when this decomposition would not make sense (as in an NP like ’Zie je nog wel eens een vlaat?’, *Do you now and then see a custards?*)?

Experiment 5.1B: Pseudowords vs. nonwords

In Experiment 5.1B, the *word* conditions were substituted by *pseudoword* conditions in order to look at the decompositional issue in more detail. These nonwords had a real word embedded within them (e.g., ’vla’ *custard* + t = ’vlaat’). As in the previous experiment, those pseudowords were compared with nonwords presented in the same sentence frames.

Method

Materials. The same preceding contexts were used as in Experiment 5.1A. This time the *real word* conditions were replaced by the *pseudoword* conditions (C) and (D).

(C: VP) Vraag jij of Jan morgen vlaat?
Are you asking whether Jan custards tomorrow?

(D: NP) Zie jij nog wel eens een vlaat?
Do you now and then see a custards?

Remember that both the pseudoword-nonword continua (e.g., *vlaat - vlaak*) and the nonword-nonword continua (e.g., *snaat - snaak*) were identical in the different types of contexts. The construction of the sentence stimuli and the continua has already been described in the Method section of Experiment 5.1A.

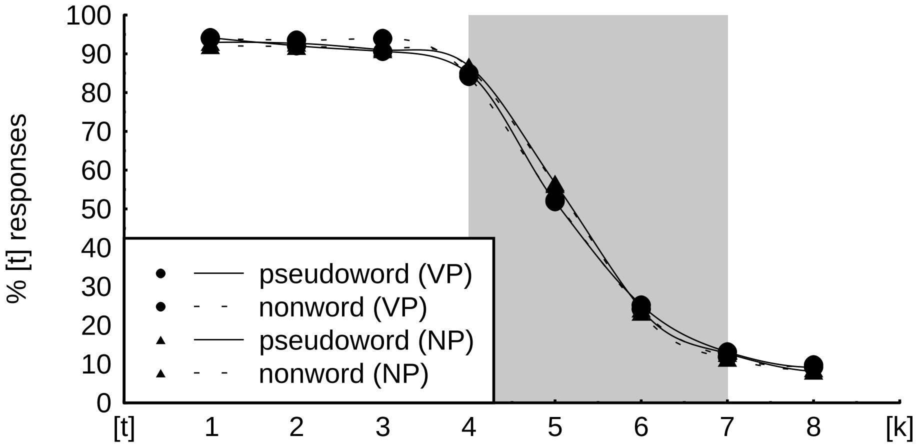


Figure 5.3: Experiment 5.1B: Proportion [t] responses (in %) for each sentence type on the [t]-[k] continuum.

Subjects. Twenty student volunteers (15 female and 5 male) from Nijmegen University were paid for their participation. They were all native Dutch speakers and none of them reported any hearing deficit.

Procedure. The procedure was identical to the one in Experiment 5.1A, as was the amount of stimuli each subject heard in each of the two sessions. The same randomized lists were used as in Experiment 5.1A with the difference that 'real word' sentences were substituted by 'pseudoword' sentences.

Results and Discussion

Again the percentage of [t] responses was computed for each subject as a function of sentence type and stimulus continuum, as plotted in Figure 5.3. The boundary region was defined as in the previous experiment from step 4 to step 7. Each listener's responses to each step in each sentence along the continuum were ranked and divided into three different RT groups. Mean RTs (measured from the onset of the final phoneme) for these groups were respectively 427 ms (SD = 102 ms), 574 ms (SD = 132 ms), and 840 ms (SD = 315 ms).

A one-way repeated-measures ANOVA on the proportion of [t] responses in the boundary region showed no lexicality effect, no context type effect and no interaction of the two factors. There were main effects of the factors item set (F(1,19) = 10.64, $p < .01$) and RT range (F(2,38) = 7.51, $p < .01$) and also the interaction of these two factors was significant (F(2,38) = 59.91, $p < .0001$).

Furthermore, the factor item set interacted with the factor lexicality (F(1,19) = 7.49, $p < .01$). The main effect of item set is not surprising since the sentence-final items contained exactly the same vowels as in the previous experiment. Therefore the same formant transitions caused an overall weaker bias towards the [t] endpoint for the monosyllabic item set (*vlaat, snaat*) in which the possible words and the nonwords contained the vowel [ɑ]. As the interaction of RT range with item set suggests, this effect was strongest in the fast RT range (F(1,19) = 69.45, $p < .0001$) but was not significant in the medium and slow RT ranges. The interaction of the factor item set with the factor lexicality was due to significantly more [t] responses to the pseudoword ’magiet’ in the verbal context (i.g., *Lach je als John haar magiet?*) than to the nonword ’meliet’ in the same context (t(19) = 3.43, $p < .01$). No other pairwise comparisons of the pseudoword condition with the nonword conditions in each item set revealed significant differences.

For each RT range one-way repeated-measures ANOVAs on the proportion of [t] responses in the boundary region were calculated. These proportions are illustrated in Figure 5.4. In the fast RT range there was a significant interaction of lexicality with context type (F(1,19) = 5.04, $p < .05$). No other factor revealed any reliable effects and none of the other interactions were significant. Although post-hoc pairwise comparisons revealed no significant effects, some of the differences were near significance: this might explain the interaction of lexicality with context type. Because none of these effects was fully reliable, they will not be discussed further. Neither in the medium nor in the slow RT range were any of the main effects significant.

In sum, no clear difference could be observed between the pseudowords and the nonwords. If people did decompose the strings ’vlaat’ and ’magiet’ into their respective subcomponents ’*noun* + t’, they did not use this decomposition to influence their categorization decisions in any systematic way (there was an effect for *magiet* but not for *vlaat*). It would appear that for there to be a reliable bias towards [t] responses, there needs to be a meaningful grammatical relationship between the two ’decomposed’ parts. The morphological status of the last phoneme in the verbal contexts on its own thus can not account for the effect found in Experiment 5.1A.

Although Experiment 5.1B suggests that pseudowords which have real words embedded within them are not treated in a systematically different way from nonwords in phonetic categorization, it is evident from advertisements, for example, that people are very well able to decompose nonsense strings into morphemes that normally do not share a grammatical relationship. Thus, when Dutch speakers read the sentence “Vlaaien we bij jullie of bij ons?” printed above the picture of a delicious looking ’vlaai’ (a special cake traditionally made in Limburg; i.e.,

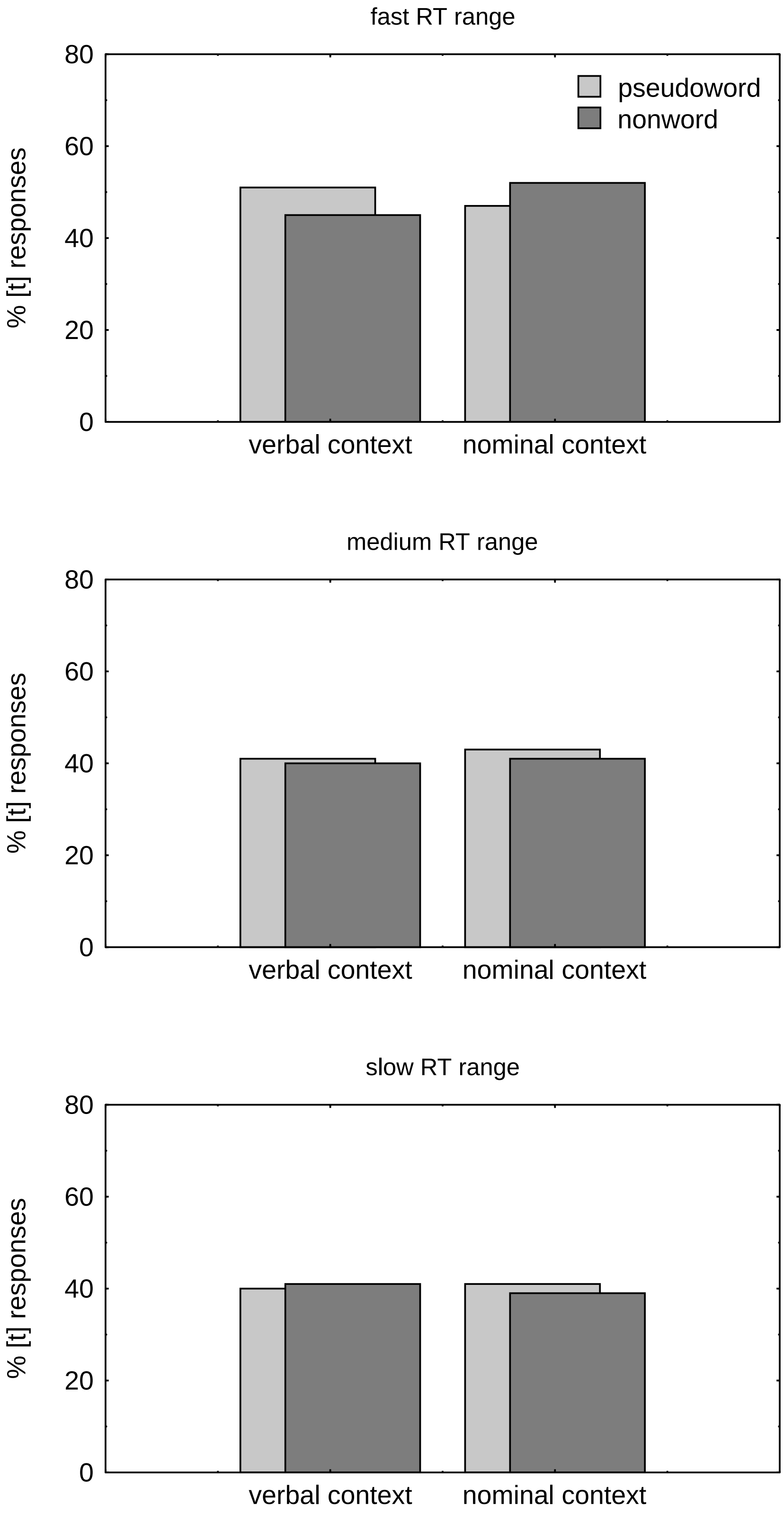


Figure 5.4: Experiment 5.1B: Proportion of [t] responses (in %) in the ambiguous continuum region for each sentence type in each RT range.

the sentence means: *Shall we “vlaai” at your place or at our place?*), they will effortlessly understand the meaning of that sentence by decomposing the string ’vlaaien’ (although not an existing Dutch word) into its two constituent parts, namely the noun ’vlaai’ and the verbal plural marker ’-en’. Similarly, they would be able to decompose the final word in ”Cum laude gevlaait” [1] into the noun part ’vlaai’ and the participle affixes ’ge-’ and ’-t’. Note, however that the latter example not only requires an appropriate situational context in order to be interpretable but also requires knowledge of the fixed expression “Cum laude geslaagt” (i.e., *Graduated cum laude*) and its meaning. In the first example, on the other hand, the sentence context might be enough to trigger the morphological parsing of the nonsense word.

Because the newly formed words in this experiment neither had an “established” status nor any semantic context information and because the morphological influence only shows up at an early stage in phonetic decision-making (as has been demonstrated in Experiment 5.1A), this ability to decompose nonsense strings might not have been of any advantage in Experiment 5.1B. It is possible that the morphological computation took too much time in order to be effective. Thus, all responses - even the slower ones - were initiated before this computation was finished or even initiated. In that sense, pseudowords were by no means different from “pure” nonwords. Future research, however, might be able to demonstrate how much exposure people need in order to establish morphological structures within novel lexical entries.

Before the results of Experiments 5.1A and 5.1B can be discussed further, one open question needs to be answered: how far are the results of Experiment 5.1 actually dependent on the processing of the preceding sentential context? Would the same effect show up if listeners were not provided with any predictive information? In order to test this question, the final words, pseudowords, and nonwords of the previous two experiments were presented in isolation in the next experiment.

If similar differences to those in Experiment 5.1A, where uninflected nouns and inflected verbs were presented in appropriate sentence contexts, are observed for words in isolation, this would indicate more general processing differences between inflected verbs as compared to uninflected nouns. It might suggest, for example, that there are morphologically decomposed representations of verbal forms. Similar lexical effects for inflected verbs and uninflected nouns, on the other hand, would support the interpretation that the effect observed for the verbal phrases in Experiment 5.1A could indeed be attributed to the influence of syntactic information on phonetic decision-making. Yet another possibility is

---

[1] Both examples were taken from advertisements of the company MULTIVLAAI.

that the difference between uninflected nouns and inflected verbs in isolation will be different from the one observed in Experiment 5.1A: there could be a larger effect for nouns than for verbs. This is not so unlikely since inflected verbs in isolation make less sense and are less common in natural speech than nouns in isolation. Without a licensing context, the lexical effect on inflected verbs might be weaker than that for uninflected nouns.

Experiment 5.2: Words, pseudowords and nonwords in isolation

Method

Materials. All eight continua that were used in sentence contexts in Experiments 5.1A and 5.1B, each with eight steps, were presented in isolation. There were two verb-nonword continua (e.g., *gaat - gaak*), two noun-nonword continua (e.g., *plaat - plaak*), two pseudoword-nonword continua (e.g., *vlaat - vlaak*), and two nonword-nonword continua (e.g., *snaat - snaak*). All sentences used in the previous experiments were created by splicing natural utterances of the preceding contexts, stored in individual sound files, onto all of the versions of the final items (words, pseudowords and nonwords) which had also been stored as individual files. Since these files already existed, no further speech editing was required. The same was true for filler items. These versions were presented to listeners without the preceding contexts. In total, there were two versions of four different continua (verbs-nonword, noun-nonword, pseudoword-nonword and nonword-nonword).

Subjects. Twenty students (15 female and 5 male) from Nijmegen University volunteered for participation. They were paid for their participation and none of them had any hearing difficulties. All of them were native speakers of Dutch. None had taken part in Experiment 5.1.

Procedure. Again four lists of stimuli were created that varied only with respect to the order of stimulus presentation. These lists were again distributed evenly across subjects. Subjects were tested in groups of one to four in a quiet room. Apart from that, the technical setup was the same as in the previous experiments. The procedure was different from the previous experiments in only two respects. First, although the number of stimulus presentations was exactly the same as in the previous experiments (768 presentations of experimental stimuli + 128 filler presentations), the inter-stimulus interval (ISI) could be reduced by

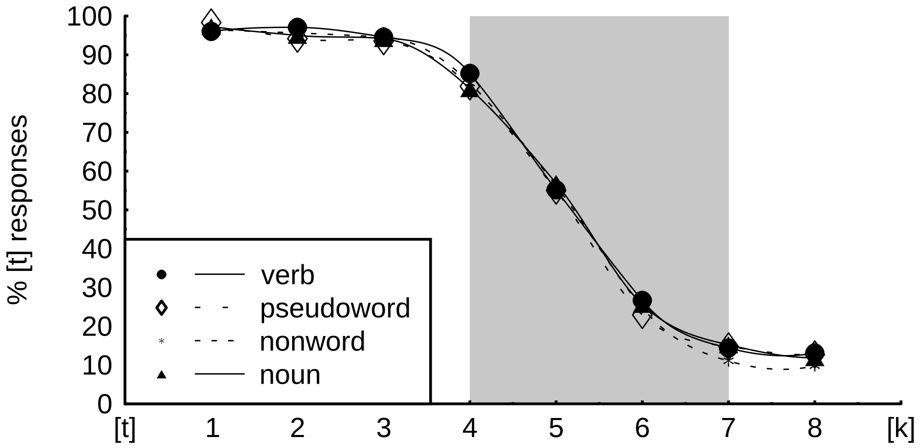


Figure 5.5: Experiment 5.2: Proportion of [t] responses (in %) for each word type (in isolation) on the [t]-[k] continuum.

one second to 1500 ms (because listeners now heard isolated words rather than whole sentences). An ISI of 2500 ms would have been much too long. Second, therefore, subjects did not have to come back a second time. After a practice session which included 24 stimuli, the main session started and was interrupted two times for breaks. The whole experiment lasted no longer than 50 minutes.

Results and Discussion

The percentage of [t] responses was computed for each subject as a function of word status (verb, noun, pseudoword, nonword) and stimulus continuum. In Figure 5.5 the categorization functions of the four different item types across item sets are plotted across subjects. As in the earlier experiments, a boundary region was defined. This extended from step 4 to step 7. Each listener's responses to each step along the stimulus continuum were ranked and divided into fast, medium and slow RT groups. The respective mean RTs (measured from the onset of the final phoneme) were 370 ms (SD = 107 ms), 516 ms (SD = 114 ms) and 739 ms (SD = 224 ms). Proportions of [t] responses were computed in the boundary region in each RT range. These are plotted in Figure 5.6.

A one-way repeated-measures ANOVA on the proportion of [t] responses in the boundary region showed no effect of word status (lexicality). There was, however, a main effect of RT range (F(2,38) = 4.72, $p < .05$) as well as an interaction of RT range both with word status (F(6,114) = 4.05, $p < .001$) and item set (whether the item was mono- or bisyllabic; F(2,38) = 17.74, $p < .0001$). Not surprisingly, the factor item set also had a significant effect on the categorization performance (F(1,19) = 13.59, $p < .01$) since the vowels [ɑ] and [i] that were used in the different item sets in the current experiment were identical to those that had already been used in Experiment 5.1. None of the other interactions was significant as were none of the post-hoc pairwise comparisons. As was also the case in previous experiments, the bias of the categorization functions towards the [t] endpoint was overall weaker for the monosyllabic item set (that contained the vowel [ɑ]) than for the bisyllabic item set (with the vowel [i]) due to different formant transitions in the two vowel contexts. This difference proved to be quite stable over time, since it was significant in both the fast (F(1,19) = 26.35, $p < .0001$) and the medium RT ranges (F(1,19) = 5.64, $p < .05$) and almost significant in the slow RT range (F(1,19) = 3.82, $p < .07$). This effect became weaker over time, as it did in Experiments 5.1A and 5.1B. This explains the interaction of the factor RT range with item set.

For each RT range one-way repeated-measures ANOVAs on the proportion of [t] responses in the boundary region were performed separately. In the fast RT range, there was a main effect of word status (whether the presented stimulus was a verb, a noun, a pseudoword or a nonword; F(3,57) = 5.29, $p < .01$) and, as already mentioned above, a significant effect of item set. The interaction of these factors was not significant. Post-hoc t-tests showed that verbs in the fast RT range received significantly more [t] responses as compared to nonwords (t(19) = 2.6, $p < .05$) while the difference between verbs and pseudowords was only marginally significant (t(19) = 1.92, $p < .07$). Nouns in the fast RT range received significantly more [t] responses than both nonwords (t(19) = 4.02, $p < .001$) and pseudowords (t(19) = 2.91, $p < .01$) but did not differ significantly from inflected verbs. Both factors were also significant in the medium RT range and again their interaction was not significant (word status: F(3,57) = 3.92, $p < .01$; item set: F(1,19) = 5.64, $p < .05$). The lexical effect that was observed in the fast RT range was reversed in the medium RT range. Nouns in the medium RT range received the fewest [t] responses as compared to the other three conditions (nouns vs. nonwords: t(19) = 3.7, $p < .01$; nouns vs. pseudowords: t(19) = 3.13, $p < .01$; nouns vs. verbs: t(19) = 2.14, $p < .05$). Although nonwords received the most [t] responses in the medium RT range, none of the other pairwise comparisons reached significance. In the slow RT range no significant effects were observed.

The reversed lexical effect in the medium RT range, in combination with the lexical effect in the fast RT range, explains the overall interaction of word status with RT range. Furthermore, this explains why there was no effect of word status in the overall ANOVA. The current data is in two respects in line with those obtained by McQueen (1991; Experiment 1) on word-final ambiguous

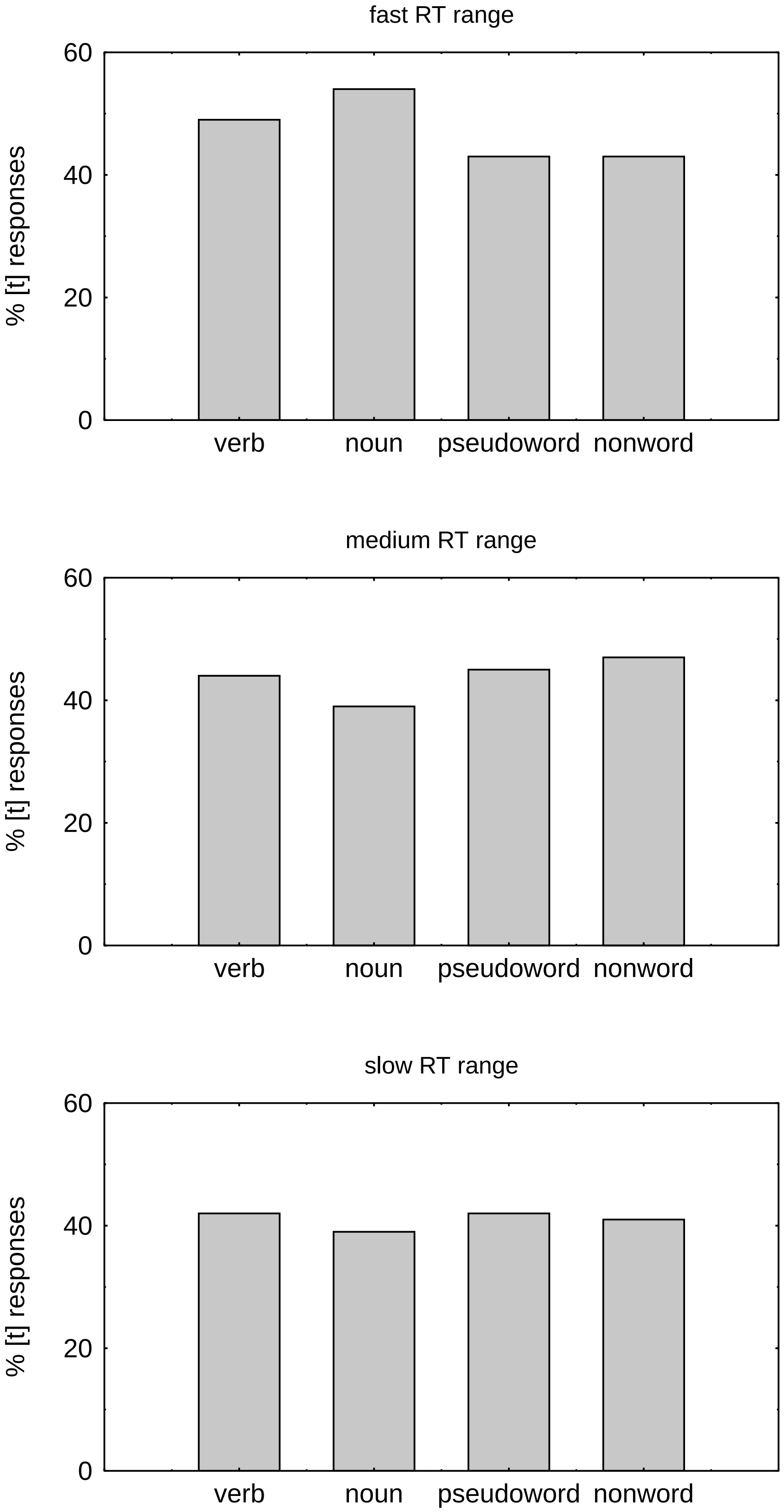


Figure 5.6: Experiment 5.2: Proportion [t] responses (in %) in the ambiguous continuum region for each word type (in isolation) in each RT range.

phonemes. McQueen (1991) also found no overall lexical shift in the categorization functions but found - just as in the present study - a reversed lexical effect in the medium RT range. Thus, in the medium RT ranges of both studies, lexically-inconsistent phonemes received more responses in the boundary region than lexically-consistent phonemes (this effect was reliable only for the nominal item set in the current study). The difference between the current data and McQueen's data, however, is that McQueen did not observe a lexical effect in the fast RT range, as was the case in the present study.

Overall, pseudowords in isolation showed similar results as the same pseudowords presented in sentence contexts. In both cases, they did not produce different categorization patterns as compared to nonwords that had no real words embedded within them. While a decomposition of those pseudowords (e.g., *vlaa* + *t*) would have been plausible in an appropriate sentence context (i.e., in a VP), this decomposition was less likely to occur in isolation where no grammatical frame triggered this decomposition.

Because of the reversal of the lexical effect (significant for nouns, not significant for verbs) in the medium RT range in the data at hand, it is hard to interpret the pattern of categorization functions in the fast RT range. It is therefore not clear whether the effect in the fast RT range is truly reliable. It would be possible to interpret the effect if it went away as RTs increased but it is hard to interpret an effect that reversed as RTs increased.

As in the McQueen (1991) study, the stimuli used here consisted of high-quality natural-sounding material. As already outlined in the Introduction, this might have prevented lexical information from exerting any influence on the categorization task because listeners could solve the task on the basis of acoustic information alone. Remember that one important question of the current study was whether there was a categorization difference between inflected verbs and uninflected nouns presented in context as compared to the same words presented in isolation. Therefore, the next experiment was conducted in which the same items were presented in a degraded version. This was based on the McQueen (1991) study, where lexical effects on word-final ambiguous sounds were only obtained when the materials were degraded by low-pass filtering. While McQueen (1991) did not observe lexical effects on high-quality words, he found a lexical effect for degraded materials that was strongest in the fast RTs. A similar pattern here would make it easier to interpret the results from Experiment 5.1A, where words were presented in sentence contexts. Materials in the current study were degraded by the addition of noise since low-pass filtering would have deleted disproportionally more cues for [t] than for [k].

Experiment 5.3: Isolated words in degraded versions

Method

Stimulus construction. The degradation of stimuli was achieved by adding noise to the target phoneme and some of the preceding vowel in each stimulus. This procedure was taken from Pitt and Samuel (1993). Each sample value of the digitized signal was changed by a random amount within a defined range (e.g., ± 200). The amplitude of the noise added to the signal could thus be varied by varying the range: the higher the range the higher the amplitude of the noise. The noise started within the final portion of the preceding vowel with a linear onset of 20 ms so that the steady state of the noise was reached at the onset of the target phoneme. The amplitude envelope also had a linear offset in the final 20 ms of the target phoneme. In order to establish the appropriate amount of noise, a pretest was conducted in which nine subjects were presented with a small subset of the items in three different noise conditions. These varied with respect to the amplitude of the noise (300, 400, 500). On the basis of the steepness of the categorization functions from the pretest, a noise range of 400 was chosen for the actual experiment.

Subjects and Procedure. Fifteen female and five male students from the subjects' pool of the Max Planck Institute for Psycholinguistics were paid for their participation. They were all Dutch native speakers and had no hearing deficits. The technical setup as well as the experimental procedure were identical to those in Experiment 5.2. None of the subjects had taken part in either Experiment 5.1 or Experiment 5.2. The experimental items were presented in the same randomized orders as in Experiment 5.2. Thus, four different lists were distributed evenly across subjects.

Results and Discussion

The results were analysed in the same way as in the earlier experiments. Figure 5.7 shows the overall categorization functions. Again the listeners' responses were divided into three RT ranges. The mean RTs (measured from the onset of the final phoneme) were 384 ms (SD = 115 ms) in the fast, 537 ms (SD = 128 ms) in the medium and 802 ms (SD = 246 ms) in the slow RT range. The respective proportions of [t] responses in the ambiguous area are plotted separately for each RT range in Figure 5.8.

The one-way repeated-measures ANOVA calculated for the boundary region (steps 4 - 7) revealed only a marginal effect of word status (F(3,57) = 2.32, p

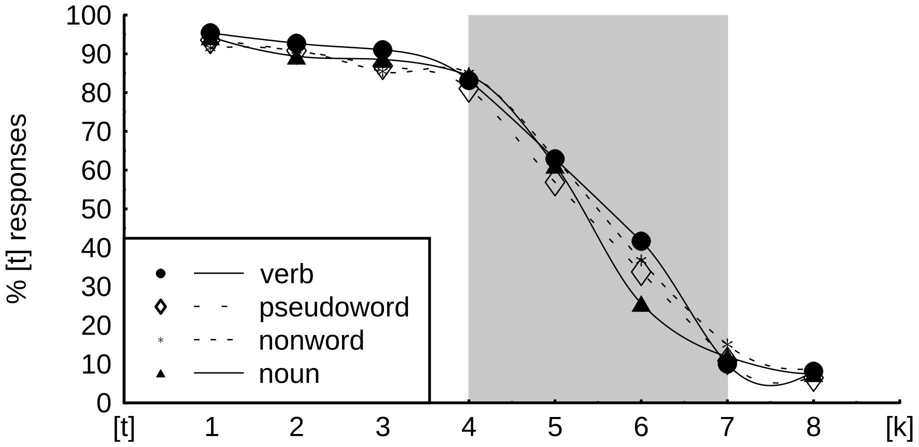


Figure 5.7: Experiment 5.3: Proportion of [t] responses (in %) for each word type (in isolation and degraded by noise) on the [t]-[k] continuum.

$<$ .08) while the effect of item set proved to be quite stable (F(1,19) = 6.3, p $<$ .05). Thus, even when the stimuli were masked by noise, formant transitions in the items containing the vowel [ɑ] were still stronger than in the items that contained the vowel [i]. There was no interaction of the factors word status and item set and also the factor RT range was not significant. There was, however, a fully reliable interaction of the factor RT range with item set (F(2,38) = 10.78, p $<$ .0001) and a marginally significant three-way interaction of RT range, item set and word status (F(6,114) = 2.13, p $<$ .06). The overall weaker [t] bias for the monosyllabic item set showed a similar time pattern as in previous experiments: the effect was strongest in the fast RT range (F(1,19) = 14.27, p $<$ .001) but was not significant in the medium and slow RT ranges. Post-hoc t-tests showed no reliable differences between the four different conditions. Only the pairwise comparisons within the bisyllabic item set revealed one significant difference: the verb ’bespiedt’ received significantly more [t] responses than the noun ’graniet’ (t(19) = 2.46, p $<$ .05).

As already mentioned above, the separate ANOVA for the fast RT range revealed a significant effect of item set and furthermore a significant interaction of item set with word status. This interaction was due to different patterns for mono- and bisyllabic items. While in the first set the pseudoword ’vlaat’ received significantly more [t] responses than the noun ’plaat’ (t(19) = 3.66, p $<$ .01), the verb ’bespiedt’ in the bisyllabic set received significantly more [t] responses than the pseudoword ’magiet’ (t(19) = 2.12, p $<$ .05). None of the pairwise comparisons collapsing across item sets revealed significant differences. ANOVAs performed

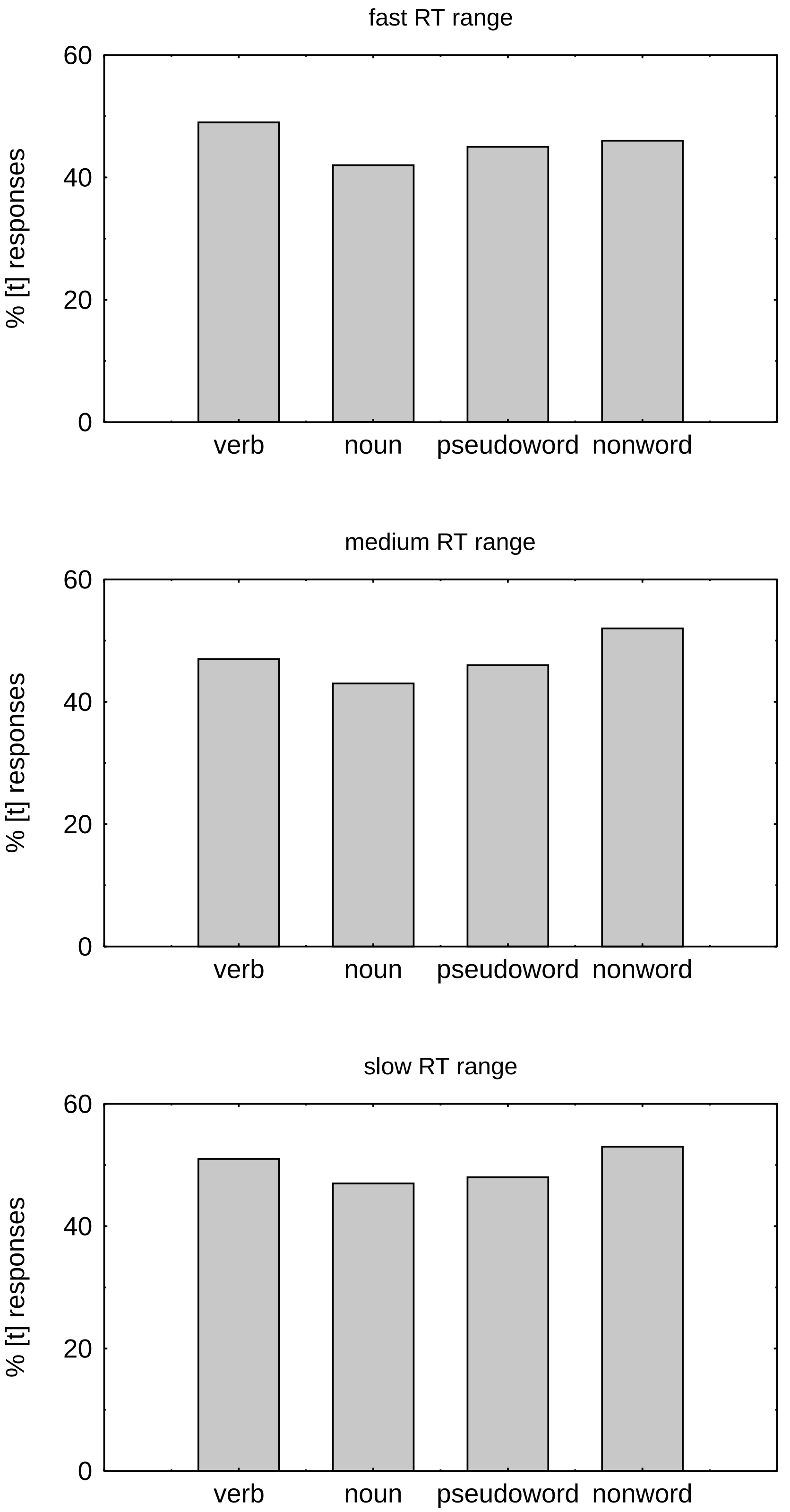


Figure 5.8: Experiment 5.3: Proportion [t] responses in the ambiguous region (in %) for each word type (in isolation and degraded by noise) in each RT range.

separately for the medium and the slow RT ranges revealed no significant effects.

In sum, although the phonemes in the degraded version were masked by noise and thus harder to hear than those in Experiment 5.2, no substantial differences in the results were obtained. The degradation was thus not severe enough to trigger lexical influences. So, maybe the results of Experiment 5.1A were true sentential effects since no lexical effects were obtained for the same items presented in isolation.

There is, however, an alternative explanation for the missing lexical effects in the last two experiments. Remember that the final words (presented in isolation in Experiments 5.2 and 5.3) were not uttered in isolation but were originally spliced from sentence contexts. Although the onsets of those words presented in sentence contexts sounded quite natural, this was not true when they were presented in isolation. The onsets sounded slightly distorted, which might have had the effect that the real words (either verbs or nouns) did not differ enough from nonwords to yield significant lexical effects. In other words, it is possible that most items sounded like nonwords when they were presented in isolation. This concern will be addressed in Chapter 6.

General Discussion

In Experiment 5.1A sentential effects were observed overall for both verbal and nominal sentence contexts. The sentence contexts biased listeners such that they gave more lexically-consistent responses to ambiguous sounds. Overall, this lexical influence was not stronger for ambiguous sounds that were in a syntactically predictable morpheme position (like the [?] in *gaa?*) than for the same sounds in a syntactically nonpredictable position (like the [?] in *straa?*). However, the time when this sentential/lexical information was most effective differed between the two conditions. The sentential effect on potential morphemes (VPs) was strongest in listeners’ fastest responses while the lexical influence on noun-final sounds was stronger in the medium RT range. When listeners initiated their phonemic decisions relatively slowly, sentence contexts did not have any influence on their decisions.

Experiments 5.1C and 5.1D examined the key question whether inflected verbs and uninflected nouns would show similar categorization patterns when they were presented in isolation as compared to those words in sentence contexts. Neither of these experiments revealed any reliable results. Overall, there were no lexical effects for verbs and nouns as compared to either pseudowords or nonwords. This was true whether the items were masked by noise or were not masked by noise. Although there were lexical effects for both inflected verbs and uninflected nouns in the fast RT range in Experiment 5.2, it is not clear what this effect really means since it was reversed as the subjects' responses got slower. One explanation for this unclear pattern might be that all isolated items were spliced out of sentence contexts so that their onsets sounded distorted. This acoustic confound probably reduced the comprehensibility of the real words so severely that they were not apprehended as any different from the nonwords.

Due to the confounds in Experiments 5.2 and 5.3 it is not possible to answer one of the two major questions of this study, namely whether the categorization functions for nouns and verbs presented in sentence contexts would be any different from those functions for the same words presented in isolation. However, the results of Experiment 5.1A do address the question whether ambiguous sounds are labeled differently as parts of inflected verbs as compared to uninflected nouns as a consequence of sentential influences.

While there were overall lexical/sentential effects in both verbal and nominal sentence contexts, the temporal pattern of these effects differed between the two different contexts. The sentential/syntactic influence on the categorization of ambiguous sounds at the ends of inflected verbs was only effective in subjects' fastest responses. This temporal pattern is in line with the results reported by Alphen and McQueen (2001) who observed syntactic effects in the fastest reactions when the relevant syntactic information had just been integrated. They argued that syntactic information has only a time-limited influence on phonemic decisions: if the relevant decision units are activated when syntactic information is being resolved, this information can bias the activation of those decision units. Generally speaking, only those decisions can be influenced by syntactic information that are close in time to the resolution (or integration) of that information.

In all experiments, the to-be-labeled sound was always the very last sound of the sentences. In the verbal condition, this last sound sometimes (if it was an unambiguous [t]) carried syntactically relevant information. Because syntactic integration occurs as soon as the final sound is reached, only those decisions that are close in time to the point of syntactic integration should be influenced by the preceding information. This is exactly what was observed: there was a lexical/sentential influence on the fastest responses when the ambiguous sound was part of a syntactically predictable verb.

Note that this result is in line with the predictions of the autonomous model Merge but not in line with the principle of feedback. Remember that TRACE predicts sentential effects to build up over time since activation from higher order levels down to lower processing levels accumulates over time. Furthermore, once sentential information has biased the phonemic/perceptual decision unit towards one or the other phoneme, this bias cannot be undone. Therefore, any sentential/syntactic effects on phonemic decisions should build up over time and should be strongest in the subjects' slowest decisions. The observed pattern of categorization functions for the verbal condition clearly contradicts this prediction.

But what about the nominal condition? In principle (as has been discussed in Chapter 4) there is no a priori reason to assume that the predictions for sentential/semantic influences on the categorization of ambiguous phonemes are different from those for sentential/syntactic influences. Thus, according to Merge, as soon as semantic integration has been completed, only those decisions that are close in time to this integration should be influenced by semantic information. In Experiment 5.1A the only reactions that were influenced by sentential/semantic information were those in the medium RT range. This might indicate that syntactic processes wrap up sooner than semantic processes and therefore both sorts of information are effective in different time windows.

But note that this implies that the integration of semantic information (in the nominal condition) had not yet started when subjects responded fast but was 'delayed' relative to the point in time when sentential/syntactic information was integrated (in the verbal condition). Was this due to the fact that the sentences were created such that the last nouns in the nominal sentence frames were semantically not highly predictable? In principle, this might be an explanation: if the final word is not strongly predicted by the semantic context, there is no a priori reason for subjects to expect a [t] at the ends of these sentences more than a [k]. The effects observed in the medium RT range for the nominal condition, therefore, might be influenced more strongly by *lexical* information rather than by *sentential/semantic* information.

In contrast to the verbal condition, where the activation of a decision unit might start relatively early on the basis of sentential/syntactic information, the activation of a decision unit in the nominal condition probably is not initiated before the onset of the final word (e.g., *straat*). Because the expectancy of a final [t] is higher (and starts earlier) in the verbal than in the nominal condition, this information might be integrated faster in the former than in the latter condition.

There are, however, two limitations to these conclusions: first, due to the high number of repetitions of each stimulus sentence, it is not clear whether these effects are really due to the processing of the sentential contexts or rather due to some strategy based on the many repetitions. Remember that each listener heard each sentence 96 times (distributed across eight continuum steps). Thus, for example, whenever they heard the first words *Vraag jij of Jan ...* they knew that the final element was either *gaat, gaa?* or *gaak* (in the real word condition) or *snaat, snaa?* or *snaak* (in the nonword condition). Therefore, not only was the final sound of the verbal phrases predictable on the basis of syntactic processing, but also the final words (or nonwords) in each sentence type were highly predictable on the basis of the experimental set-up. The effects observed in Experiment 5.1A might therefore reflect to a certain extent high predictability rather than sentential/syntactic or semantic parsing. If this were true, however, why then were the results different for inflected verbs and uninflected nouns (as was observed in Experiment 5.1A)? This difference might be attributable to the greater predictability of [t] in verbal than in nominal contexts.

The second limitation concerns the frequency issue: since Experiments 5.2 and 5.3 did not produce interpretable results, it is not clear how far the results obtained in Experiment 5.1 might have been influenced by the factor frequency. Remember that the verbs were more frequent than the nouns so that the different effects for inflected verbs and uninflected nouns in sentence contexts might be attributable to the frequency difference.

Chapter 6, therefore, reports new experiments which sought to disentangle the factor syntactic/semantic predictability from the factor experimental predictability. New materials were created that had more variability in both the sentence contexts and the final words. In order to deal with the acoustic confound in the isolated words of Experiments 5.2 and 5.3, the new final words were recorded in sentence contexts and in isolation so that their onsets would not be distorted due to speech editing. Furthermore, an attempt was made to balance the mean surface frequency of the final words between inflected verbs and uninflected nouns.

When real words were replaced by *pseudowords* (Experiment 5.1B), listeners' phonemic decisions were not influenced by sentential information. This suggests that the potential decomposition into two meaningful units that otherwise do not have a grammatical relationship does not influence listeners phonetic categorization performance. However, a new experimental set-up might be able to reveal sentential effects on pseudowords. In such a new set-up, listeners could, for example, be provided with a short story that introduced the pseudowords in appropriate contexts so that listeners would get the opportunity to *learn* the novel words before they were tested on those novel words in categorization.

The role of morphology in phonetic decision-making: Part III

Chapter 6

Although Experiment 5.1A indicated that inflectional morphology might play a special role in the categorization of ambiguous sounds, the confound of many repetitions of the stimuli constrained the conclusions which could be drawn. The next question is therefore this: will the difference between the categorization of ambiguous sounds as parts of inflected verbs and the categorization of the same sounds as parts of noun stems also hold up when listeners are presented with more variable experimental stimuli? And furthermore: was the lack of a lexical effect in Experiments 5.2 and 5.3 (where words were presented in isolation) due to acoustic confounds which were the consequence of speech editing procedures, or did this pattern indicate that the effects observed in Experiment 5.1A were purely sentential in origin?

If the confound of repetition did not influence the categorization of ambiguous sounds in Experiment 5.1A, a similar pattern of results should be obtained with more variable materials: there should be overall sentential/lexical effects for inflected verbs and uninflected nouns in sentence contexts. But the time pattern of the effects should differ such that the sentential/syntactic effect for verbs should be strongest in listeners' fast responses and should become weaker over time, while the sentential/semantic effect for nouns should be strongest in the medium RT range, as was observed earlier.

If the high proportion of stimulus repetitions did influence subjects' categorization strategies in Experiment 5.1A, however, different results might be obtained when the final words are less predictable than they were in Experiment 5.1A. It might be the case that the strong sentential/syntactic effect observed in the fast RT range for verbal phrases in Experiment 5.1A was driven by the sentential/syntactic expectancies of a final morpheme combined with the strong experimental expectancies of a certain lexical element (i.e., either *gaa?* or *bespied?*). More variable verbal phrases still predict inflectional morphemes at the ends of the final words (e.g., the [t] in *Vraag jij of Jan morgen gaat?*) but do not predict the whole word to the same extent. Of course, high predictability was also a problem for the nominal sentences in Experiment 5.1A: although the final sound was not grammatically predictable in noun phrases, the final nouns were experimentally as predictable as the final verbs. It was therefore possible that an experiment with more variability could disentangle syntactic from experimental predictability and thus also disentangle the influences of these factors on phonetic categorization.

A new set of stimuli was therefore designed that contained more experimental sentences. More stimulus variation was now possible because the *pseudoword* condition which had real words embedded within them (e.g., *vla*, ’custard’, in *vlaa* + *t*) - which had imposed high constraints on the stimulus design - was not included. As already mentioned in the previous chapter, a different experimental set-up might be needed in order to test perception of newly-created words in sentence contexts. In Experiment 6.1 inflected verbs and uninflected nouns were presented in appropriate sentence contexts while these final words were presented in isolation in Experiment 6.2. In order to avoid the potential acoustic confound - distorted word onsets generated by speech-editing - that might have been responsible for the missing lexical effects in Experiments 5.2 and 5.3, these final words were also recorded in isolation so that their onsets remained intact.

Experiment 6.1

Method

Materials. As in Experiment 5.1 experimental sentences were again questions that either ended with a verb or a noun. Also in parallel to that experiment, each [t]-final item within each set was paired with a [k]-final nonword to form three different continua: verb-nonword (e.g., *gaat* - *gaak*), noun-nonword (e.g., *plaat* - *plaak*), and nonword-nonword (e.g., *snaat* - *snaak*).

The sentence contexts that were used in Experiment 5.1 were supplemented by four new sentence frames in both the verbal and the nominal conditions, yielding a total of five different VPs and five different NPs. As in Experiment 5.1A, the contexts were constructed such that no word included either a [t] or a [k] in order to avoid any contrast effects on the categorization of the final sound. Each of these sentence frames ended with two different word-nonword and two different nonword-nonword continua, so that the experimental expectancy of the final items was lower than in Experiment 5.1. For each final inflected verb, a phonologically-matched noun appeared in the nominal condition (like ’gaat’ and

’plaat’ in the previous experiments) and phonologically-matched nonwords were presented in both the corresponding verbal and nominal sentence frames. The mean surface frequency of the inflected verbs was 203 (per 42 million based on the CELEX computerised database of Dutch) while the mean surface frequency of the uninflected nouns was 503 (per 42 million). While each verb-nonword continuum and each noun-nonword continuum only occurred in the appropriate sentence contexts, the matched nonword-nonword continua occurred in both sentence frames. The phonological matches were achieved in the following way: the offsets of words and nonwords always matched from the last vowel on. Three different word offsets occurred in the stimuli: long V + t (e.g., *gaat* ’leaves’, *plaat* ’record’), short V + t (e.g., *bezet* ’occupies’, *buffet* ’buffet’), and short V + s + t (e.g., *wast* ’washes’, *gast* ’guest’). This procedure yielded a total of 40 different sentences and 30 [t]-[k] continua (10 verb-nonword, 10 noun-nonword, and 10 nonword-nonword continua). Note that each nonword-nonword continuum occurred twice (once in a verbal and once in a nominal context), while each verb-nonword and noun-nonword continuum occurred only once, which explains the discrepancy between the total number of continua and the total number of experimental sentences. The complete set of items is listed in Appendix E.

Stimulus construction. The sentences were recorded by the same female speaker who had also spoken the materials for the previous categorization experiments. The technical recording procedure was exactly as in the previous experiments. Again a fifteen-step place-of-articulation continuum was constructed using the same procedure as in earlier experiments (Repp, 1981; Stevenson, 1979). The [k] was spliced out of a natural waveform from the nonword ’bandiek’ that matched the real noun ’bandiet’ (*bandit*) and the [t] was spliced out of a natural waveform from the nonword ’mediet’ from the nonword-nonword continuum (*mediet - mediek*) that was matched phonologically to the noun-nonword continuum *bandiet - bandiek*. Both tokens were spliced such that their durations were identical (250 ms). The vowels or VC clusters that preceded the final target phoneme were kept constant within the matched verb, noun and nonword triplets (e.g., *gaak, plaak, snaak*). The transitions in those preceding phonemes again cued [k] so that any bias based on these transitions would work against the predictions. The sentence frames before the onsets of the final words were kept as they had been uttered by the speaker so that splices only occurred before and within the final words and nonwords. The fifteen tokens from the [t]-[k] continuum were then spliced onto all contexts.

A pretest was conducted to establish the ambiguous region of the fifteen-step continuum. Eight native Dutch listeners were presented with a subset of the whole range of experimental stimuli and were asked to label the sentence-final sounds as either [t] or [k]. In this subset, two noun-nonword continua were presented in the appropriate sentence contexts while the matched nonword-nonword continua were presented in the verbal contexts. Two different item sets were presented in the opposite order: while the verb-nonword continua were presented in the appropriate verbal contexts, the matched nonword-nonword continua were presented in the nominal contexts. One other item set was presented in yet another order: both the verb-nonword continuum and the matched noun-nonword continuum were presented in the appropriate sentence contexts. Thus, 10 different sentences were presented in the pretest. Each of the fifteen steps was presented twice in each of the 10 sentences so that each listener heard a total of 300 stimuli. Visual inspection of the categorization functions showed that steps 1 - 4 and 11 - 15 were clear instantiations of a [t] and a [k] respectively, while steps 5 - 10 were ambiguous. Steps 4 - 11 were thus chosen for presentation in the experiment and will henceforth be called steps 1 - 8.

Subjects. Twenty students (16 female and 4 male) from Nijmegen University participated in the experiment and were paid for each session they attended. They were all native Dutch speakers and had no known hearing problems. None of them had participated in any experiment reported in Chapter 5.

Procedure. Four differently randomized lists were created that contained all of the stimuli while only the order of stimulus presentation varied between the lists. The lists were distributed evenly across the participants. The subjects were tested in groups of one to four in a quiet room. The technical facilities, the instructions and the experimental setup were the same as in Experiments 5.1 (i.e., where also whole sentences had been presented). Each of the eight continuum-steps of each of the forty sentences was presented three times to each subject which summed up to a total of 960 stimulus presentations for each participant. The ISI was 2500 ms. As in Experiments 5.1A and 5.1B subjects had to come back for a second session within one week because of the large number of stimuli. Each of the sessions started with a practice block of 24 sentences before the actual experiment. In each session, experimental blocks were interrupted twice for breaks. Each session contained 480 stimulus presentations and lasted about 50 minutes.

Results and Discussion

The percentage of [t] responses was computed for each subject as a function of sentence type and stimulus continuum. The categorization functions of the four

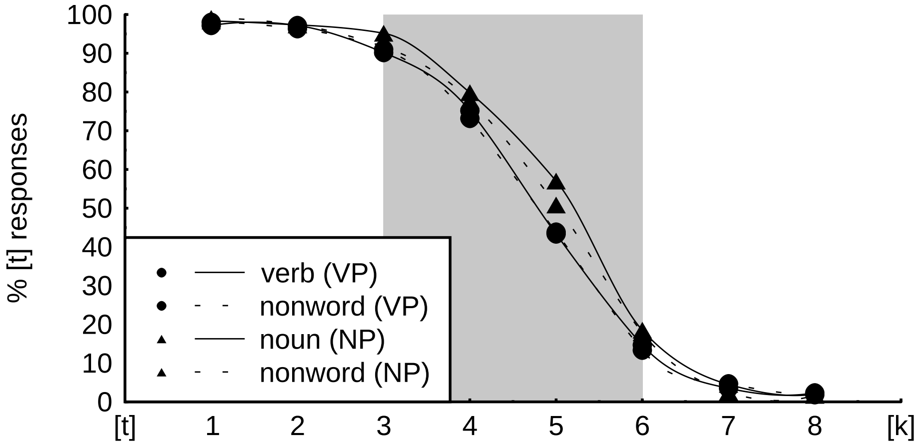


Figure 6.1: Experiment 6.1: Proportion of [t] responses (in %) for each sentence type for each step of the [t]-[k] continuum.

different sentence types are plotted in Figure 6.1 across item sets and subjects. The boundary region extended from step 3 to 6. As in the previous experiments, all analyses were computed on the proportion of [t] responses in this ambiguous continuum region. Each listener's responses to each step of the continuum in each sentence were ranked and divided into three RT groups. The mean RTs (measured from the onset of the final phoneme) were 436 ms (SD = 121 ms) in the fast, 585 ms (SD = 158 ms) in the medium and 860 ms (SD = 310 ms) in the slow RT range. In each of the three RT ranges, the percentage of [t] responses to each stimulus were computed for each subject as a function of context bias, as plotted in Figure 6.2.

In a one-way repeated-measures ANOVA there was a highly significant main effect of the factor sentence type (whether the context predicted a verb or a noun; F(1,19) = 22.63, $p < .0001$) while the factor lexicality (whether the final element was an existing word or a nonword) was only marginally significant (F(1,19) = 3.01, $p < .1$). There was no interaction of these factors. As in all other experiments, the factor RT range (whether subjects' responses were ranked fast, medium or slow) was significant (F(2,38) = 3.23, $p < .05$) as well as its interaction with the factor sentence type (F(2,38) = 4.51, $p < .05$). Also the three-way interaction of RT range, sentence type and lexicality was significant (F(2,38) = 3.52, $p < .05$).

A post-hoc t-test showed that nouns received significantly more [t] responses than their nonword counterparts in the same sentence contexts (t(19) = 2.16, $p < .05$) while there was no difference in the proportion of [t] responses for verbs as compared to their nonword controls (t(19) = .51, n.s.). Because nonwords presented in the nominal condition not only received more [t] responses than the same nonwords presented in verbal contexts (t(19) = 4.16, $p < .001$) but also received more [t] responses than *real* verbs (t(19) = 2.18, $p < .05$), the factor lexicality was only marginally significant in the overall ANOVA. For the same reason the interaction of lexicality with sentence context was not significant.

For each RT range, separate one-way repeated-measures ANOVAs were performed on the proportion of [t] responses in the boundary region. In the fast RT range there were main effects of both the factor lexicality (F(1,19) = 15.19, $p < .001$) and the factor sentence type (F(1,19) = 15.72, $p < .001$) and their interaction was also significant (F(1,19) = 7.63, $p < .01$). Post-hoc t-tests showed a very clear pattern: nouns received significantly more [t] responses as compared to their nonword controls (t(19) = 4.36, $p < .0001$), while this lexical effect was not observed for real verbs and the respective nonword controls. A t-test that took items rather than subjects as the repeated measure further supported this pattern of results. There were significantly more [t] responses for nouns than for their nonword controls (t(9) = 4.06, $p < .01$), while there was no such difference for the verbal condition. This categorization pattern explains the interaction of the factor lexicality with the factor sentence type.

In the medium RT range the two main effects were still significant while the interaction was not (lexicality: F(1,19) = 9.02, $p < .01$; sentence type: F(1,19) = 13.35, $p < .01$). The sentence type effect can be explained by an overall bias towards the [t] endpoint for the nominal contexts as compared to the verbal contexts. Thus, nouns got significantly more [t] responses than verbs (t(19) = 2.72, $p < .01$) and nonwords in the nominal condition received significantly more [t] responses than the same nonwords in verbal sentence contexts (t(19) = 2.23, $p < .05$). Note that only the latter pairwise comparison is fully justified because the nonwords are the same across sentence types while the real verb and real noun conditions vary with respect to both the final items and the sentence frames. This comparison is only reported in order to explain the overall effect of sentence type. The significant influence of the factor lexicality is due to a reversed lexical effect that was present for both conditions but significant only for the verbal condition. There were thus more [t] responses for the verbal nonword than the real verb (t(19) = 2.33, $p < .05$). No significant effects were observed for the slow RT range.

The results of Experiment 6.1 are very different from those obtained with less stimulus variability in Experiment 5.1A: while there was a strong lexical effect for the verbal condition in the fast RT range of Experiment 5.1A, there was absolutely no lexical effect in the fast responses to the verbal phrases of Experiment

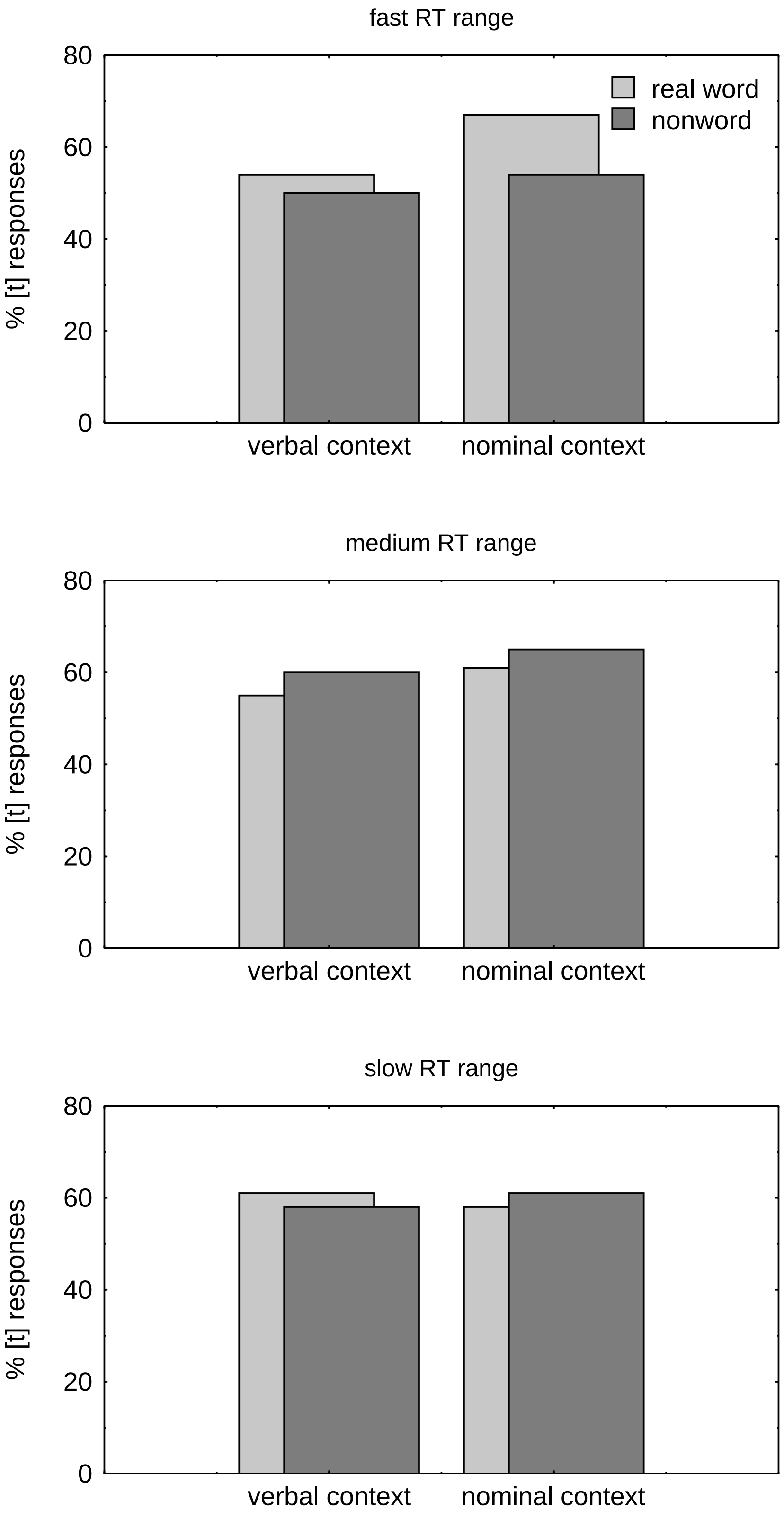


Figure 6.2: Experiment 6.1: Proportion of [t] responses (in %) for each sentence type for the ambiguous continuum region in each RT range.

6.1. Furthermore, while the lexical effect for the nominal condition was observed in the medium RT range of Experiment 5.1A, the lexical effect for the same condition in Experiment 6.1 was only obtained in the fast RT range. What was consistent across the two experiments, however, was that in both cases different patterns were observed for verbal phrases as compared to nominal phrases.

Three major questions need to be answered: first, why was there no lexical effect for VPs in Experiment 6.1? Second, why did the noun phrases show different time patterns across Experiments 5.1A and 6.1? And third, why did the two experiments produce such different results?

The major difference between the two experiments was the larger variability of the experimental sentences in Experiment 6.1 as compared to Experiment 5.1A. It was already hypothesised in Chapter 5 that the high predictability of the final items in Experiment 5.1A might have been problematic. It might be that the strong lexical effect observed in the fast reactions given to the verbal condition was mainly driven by the high *experimental* predictability rather than by *syntactic* predictability alone. Since there was also a strong *experimental* predictability for nouns in Experiment 5.1A but no *syntactic* predictability, the lexical effect may have been restricted to the medium RT range for NPs. This, however, does not explain why the lexical effect for NPs was in the fast RT range of Experiment 6.1 where nouns were less predictable than in Experiment 5.1A.

It is possible that the null effect observed for VPs in Experiment 6.1 might be due to morphological parsing: although the final inflected verbs were not as predictable as in Experiment 5.1A, the final morpheme was still predictable on the basis of syntactic information. As soon as the subjects heard the onset of a sentence like ’Vraag jij of Jan morgen ...’, for example, they could not predict the final verb but knew that the final sound was in a morpheme position: therefore, they might have been able to segment the final sound (via morphological decomposition) from the rest of the sentence and thus might have been able to base their decisions purely on phonemic information rather than on combined phonemic/lexical/sentential information.

Since the last phoneme in an NP like ’Zie jij nog wel eens een ...’ is not in a morpheme position, this decompositional strategy could not be used for the categorization of phonemes that were part of noun stems. If the null effect for the verbal condition were indeed due to morphological parsing triggered by sentential information, a different pattern of results should be obtained for the same inflected verbs presented in isolation. Since no preceding context would motivate a decompositional strategy as described above, listeners should not be able to treat the final sound of a word like ’gaat’ in isolation from its context.

Note that the effects for nouns (in Experiments 5.1A and 6.1) are very similar to effects that have been observed in other studies that investigated nouns in isolation (e.g., McQueen, 1991 and Pitt & Samuel, 1993). This suggests that the effects for nouns in the current study might be due to lexical rather than sentential information. If the effects for nouns observed in both Experiments 5.1A and 6.1 are indeed driven mainly by lexical rather than by sentential information, the pattern of results for uninflected nouns in isolation should be similar to those for the same words in sentence contexts.

Experiment 6.2: Words in isolation

The final words and nonwords from Experiment 6.1 were presented in isolation. Would the same pattern of results emerge or would there be lexical effects for both inflected verbs and uninflected nouns?

Method

Materials and Stimulus construction. The experimental stimuli consisted of the 30 continua (i.e., 10 verb-nonword, 10 noun-nonword, and 10 nonword-nonword continua) that had formed the final elements of the experimental sentences in Experiment 6.1. Since all items were presented in isolation, the nonword-nonword continua as well as verb-nonword and noun-nonword continua were only presented once. Because no interpretable effects were obtained in Experiment 5.2 where final words were spliced out of sentence contexts and presented in isolation, this time the final words and nonwords were recorded in isolation. As already argued earlier, the overall null results of Experiments 5.2 and 5.3 might have been due to the distorted word onsets that were created by excising the words from continuous speech. In order to avoid this acoustic confound, the same female speaker recorded all final words from Experiment 6.1 without the preceding sentences. The technical recording procedure was identical to the previous experiments.

A new fifteen-step place-of-articulation continuum was constructed using the same procedure (Stevenson, 1979; Repp, 1981). The [t] was spliced from the natural waveform of the verb ’praat’ (*talks*) and the [k] was spliced from the natural waveform of the nonword ’mejaak’ and the cuts were made at zero crossings. Both tokens were spliced such that their durations were identical (185 ms). Further stimulus construction followed the same principles as in Experiment 6.1. In a pilot study, a subset of 8 continua (2 verb-nonword, 3 noun-nonword, and 3 nonword-nonword continua) was presented to eight listeners in order to determine the ambiguous area of the continuum. Each step of each continuum was presented twice to each listener which summed up to a total of 240 presentations in the pretest. Visual inspection of the categorization functions revealed that steps 1 - 4 of the continuum were still clear instantiations of [t], while steps 11 - 15 were clear [k] sounds. The six intermediate steps were ambiguous. Steps 4 - 11 were thus chosen for presentation in the experiment and will henceforth be called steps 1 - 8.

Subjects. Twenty students (17 female and 3 male) from Nijmegen University were paid for their participation in the experiment. All were native speakers of Dutch and none of them reported any hearing impairments. They had not participated in any of the previous experiments reported in Chapter 5 and 6.

Procedure. Four experimental lists were created that contained all of the stimuli while only the randomized order of stimulus presentation varied between the lists. Each list was presented to 5 subjects. Subjects were tested in groups of one to four in a quiet room. The technical facilities, the instructions and the experimental setup were the same as in the previous experiment with two exceptions. Because the number of stimuli was reduced from 40 sentences in the previous experiment to 30 words in the current experiment (nonwords had to be presented only once), the total number of stimuli was smaller: each step of each continuum was presented three times which summed up to a total of 720 stimulus presentations (as compared to 960 in the previous experiment). The ISI was also reduced from 2500 ms to 1500 ms because words were presented rather than whole sentences. Therefore, all stimuli could be presented in one experimental session. Subjects heard a practice block of 24 words before the actual experiment started. They were allowed two breaks in between and each session lasted about 30 minutes including breaks.

Results and Discussion

For each subject the percentage of [t] responses was computed as a function of word type and stimulus continuum. The respective categorization functions are plotted across subjects in Figure 6.3. Each listener’s reactions to each step in each word condition along the continuum were ranked and divided into fast, medium and slow RT groups. The mean RTs (measured from the onset of the final phoneme) for each of these ranges were respectively 392 ms (SD = 85 ms), 517 ms (SD = 85 ms), and 697 ms (SD = 157 ms). The proportion of [t] responses in the ambiguous region are plotted separately for each of these RT ranges in Figure 6.4.

The boundary region was defined as extending from step 3 to step 6. A one-

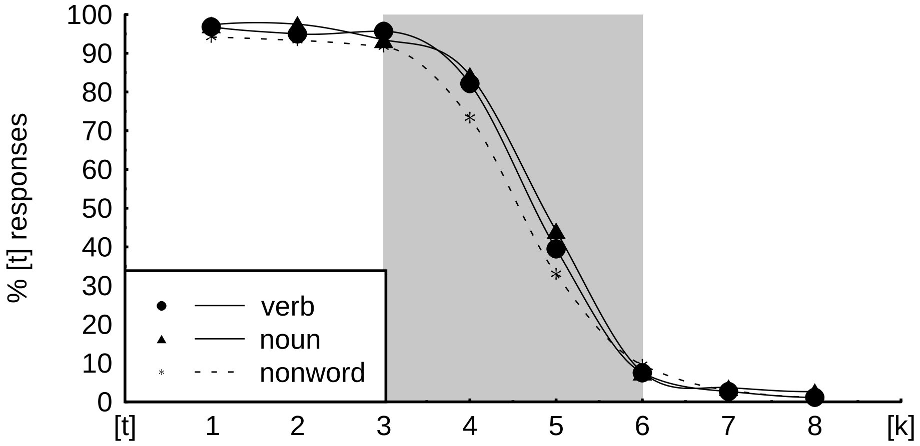


Figure 6.3: Experiment 6.2: Proportion [t] responses (in %) for each word type for each step of the [t]-[k] continuum.

way repeated-measures ANOVA on the percentage of [t] responses in the ambiguous area showed a significant effect of word status (verb, noun, or nonword; F(2,38) = 15.86, $p < .0001$) and of RT range (whether subjects' responses were ranked fast, medium or slow; F(2,38) = 5.27, $p < .01$). The interaction of these factors was also significant (F(4,76) = 9.15, $p < .0001$). A post-hoc t-test revealed that both inflected verbs and uninflected nouns received significantly more [t] responses than the matched nonwords (verbs vs. nonwords: t(19) = 4.2, $p < .0001$; nouns vs. nonwords: t(19) = 4.93, $p < .0001$). While the lexical effect for nouns vs. nonwords was also significant when a post-hoc t-test took items as the repeated measure (nouns vs. nonwords: t(9) = 3.43, $p < .01$), this was not the case for the verb-nonword comparison.

Separate one-way repeated-measures ANOVAs were then performed for each of the RT ranges. In the fast RT range, there was a strong main effect of the factor word status (F(2,38) = 21.55, $p < .0001$). This effect was stable since it was also significant in an ANOVA that took items as the repeated measure (F(2,18) = 6.49, $p < .01$). Post-hoc t-tests revealed strong lexical effects for both verbs and nouns as compared to the nonword controls (nouns: t(19) = 5.65, $p < .0001$; verbs: t(19) = 4.74, $p < .0001$). While this lexical effect for nouns was also reliable in the items analysis (t(9) = 6.18, $p < .0001$), it was only marginally significant for the verbs (t(9) = 2.18, $p < .06$).

There was no effect of word status in either the medium or the slow RT range. However, this factor was marginally significant in the slow RT range (F(2,38) = 3.16, $p < .06$). As a post-hoc t-test revealed, this effect was due to a reversed lex-

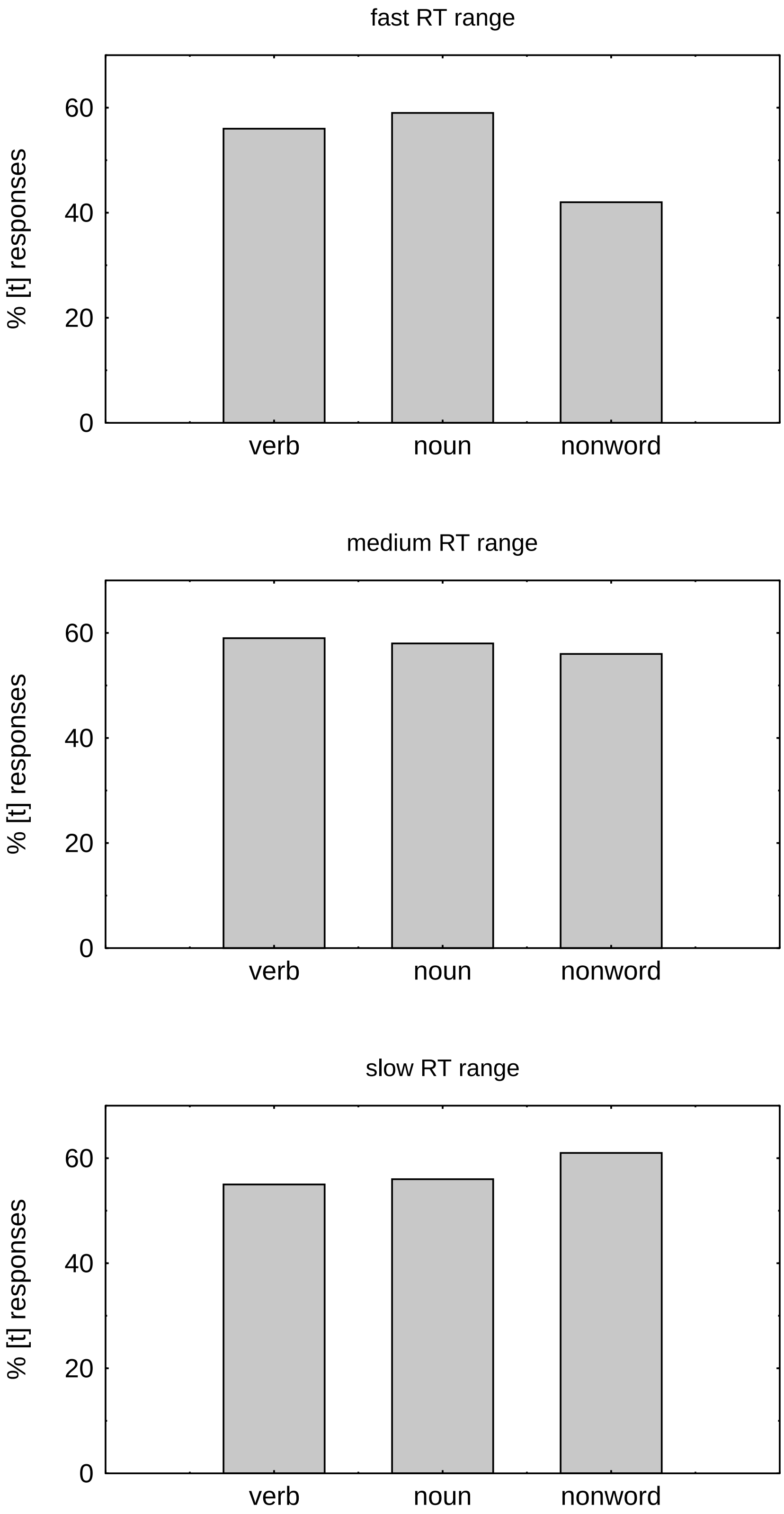


Figure 6.4: Experiment 6.2: Proportion [t] responses (in %) for each word type for the ambiguous region of the [t]-[k] continuum in each RT range.

ical effect for the verbal condition: nonwords got significantly more [t] responses than verbs (t(19) = 2.63, $p < .05$).

In contrast to Experiment 6.1 (where there was no sentential/lexical effect for VPs at all) there was a clear lexical effect for inflected verbs that was due to subjects' fastest responses. The pattern observed for uninflected nouns presented in isolation were identical to those obtained when the same nouns were presented in sentence contexts: there was a strong bias towards [t] responses in real nouns in the fast RT range irrespective of whether these nouns were presented in appropriate sentence contexts or in isolation. Furthermore, this lexical effect was more robust for uninflected nouns in isolation than for inflected verbs in isolation. While nouns thus produced almost identical results when they were presented in sentence contexts as when they were presented in isolation, inflected verbs produced very different results dependent on whether they were presented in appropriate sentence frames or whether they were presented in isolation. The implications of these results will be discussed in more detail in the context of the results reported in Chapters 4 and 5.

General Discussion

The current series of experiments tested whether sentential context would bias the categorization of ambiguous sounds differently when they were potential morphemes of inflected verbs as compared to phonemes as parts of noun stems. Experiment 4.1 was run as a pilot study in order to establish whether the categorization of sentence-final ambiguous sounds would produce interpretable effects since some necessary manipulations had never been tested before. Minimal phrases like 'de tante gaat' (VP) or 'een brede straat' (NP) were presented that could end in real words (either inflected verbs or uninflected nouns), pseudowords (real words embedded in nonwords), or nonwords. Due to some confounds in the experimental stimuli it was not possible to draw any strong conclusions about differential effects for inflected verbs as compared to uninflected nouns. Nevertheless, there was an indication of sentential effects on the categorization of ambiguous phonemes in pseudowords.

In Experiments 5.1A and 5.1B full sentences were used rather than minimal phrases. The former experiment looked at real verbs and nouns in contexts while the latter examined pseudowords (e.g., *vlaat*) in the same contexts. No sentential/lexical effects were obtained for pseudowords (Experiment 5.1B). The decomposition of morphologically complex strings therefore seems to depend on a grammatical relationship between the morphemes that needs be established better than was the case in Experiment 5.1B. Listeners probably need some sort of contextual information that triggers the decomposition not only syntactically but also semantically. Moreover, listeners may need more exposure to novel words in order to obtain decompositional effects in phonetic categorization.

There were, however, sentential or lexical effects for real verbs and nouns (Experiment 5.1A). Interestingly, although similar overall, these effects showed different patterns over time: while there was a sentential/lexical effect in the categorization of ambiguous sounds presented in VPs in the fast RT range, a lexical effect for the same ambiguous sounds in NPs was only observed in the medium RT range. The result for VPs was interpreted as a morpho-syntactic effect on subjects' fastest categorization responses. Furthermore, it was argued that the time pattern of those results was in line with predictions of the autonomous model Merge but not with predictions made by interactive models like TRACE. The slower effect in the NP condition was attributed to the fact that there were no sentential biases in NPs like those in the VP condition which strongly predicted the final sound. Since the biasing activation for [t] could have started earlier in VPs than in NPs, the effect for NPs was somewhat slower.

This interpretation was further tested in Experiments 5.2 and 5.3 where the same words were presented in isolation and where the question was whether similar result patterns would be obtained. Neither of these experiments revealed any robust lexical effects. Two possible explanations were provided: either the effects obtained in Experiment 5.1A were pure sentential effects or the null effects of Experiments 5.2 and 5.3 were due to an acoustic confound. The speech editing procedure might have distorted the onsets of the items so that - when presented in isolation - the real words might have sounded like nonwords.

It was thus not possible to interpret the results of Experiment 5.1A unambiguously. Were the effects observed there indeed sentential or would the same differences show up with words in isolation that were of a better acoustical quality? Furthermore, there was the confound of the high proportion of repetitions in the experiments of Chapter 5. It was therefore not clear whether the effects obtained in Experiment 5.1A were due to the processing of the context or rather to the high predictability of the experimental sentences.

Thus, in Experiment 6.1, the variation in the experimental stimuli was increased in order to decrease the predictability of the final words. In contrast to the results of Experiment 5.1A, there were overall more [t] responses for the *nominal* sentence contexts as compared to *verbal* sentence contexts. This was true irrespective of whether the final element was a real noun or a nonword. This stronger [t] bias for NPs was quite robust since it was also significant in the items analysis. Neither in the overall analysis nor in the different RT ranges was there a lexical effect for VPs. Even stranger, this null effect became a reversed lexical effect in the medium RT range where nonwords in verbal sentence contexts received more [t] responses than real verbs.

In contrast to Experiment 5.2 (where words were also presented in isolation), there were robust lexical effects for both inflected verbs and uninflected nouns when tokens of these words which had been recorded in isolation were presented without preceding contexts (Experiment 6.2). This overall effect was due to subjects' fastest responses since no lexical effects were observed in the two slower RT ranges.

The results from Experiment 6.2 are in line with those obtained in earlier studies (e.g., McQueen, 1991; Pitt & Samuel, 1993). Both McQueen (1991) and Pitt and Samuel (1993) reported lexical influences on the identification of word-final ambiguous nouns, although McQueen (1991) observed these influences only when his stimuli were degraded. Moreover, in both studies, lexical influences on the categorization performance were strongest in subjects' fastest responses and vanished or decreased over time. The same pattern was observed in Experiment 6.2. Importantly, there were no differences in the categorization functions for uninflected nouns in isolation as compared to inflected verbs in isolation with one exception: the effects observed for uninflected nouns were more robust as compared to the effects obtained for verbs. Since the mean surface frequency for uninflected nouns was slightly higher than the one for inflected verbs, the weaker effect for verbs might be a reflection of that frequency difference. Note, however, that this difference in frequency cannot fully account for the complete absence of a lexical effect for inflected verbs in sentence contexts. If frequency had played a major role in listeners' phonetic decisions irrespective of whether words were presented in contexts or in isolation, no different pattern should have been observed for inflected verbs in coherent sentence frames as compared to the same verbs presented in isolation.

As to the feedback debate, the results from Experiments 6.1 and 6.2 are in line with the predictions of Merge but contradict the predictions based on TRACE. While Merge assumes that sentential/lexical information should exert the strongest influences in listeners' fast responses, TRACE predicts that these influences should be stable over time or - if anything - build up over time. In the current study, the lexical effect for both verbs and nouns was strongest in the fast RT range and vanished as the responses become slower. Merge can account for this time-course of the effect: lexical as well as phonetic activation feeds information forward to a decision level. This information can have an effect (very much like sentential information) as long as not all information has been fully integrated. As soon as a word has been accessed, form information (that can bias phonetic decision) is no longer necessary. If a decision is initiated too long after the point in time when the word has been accessed, this form information might not be available anymore for the decision.

What further implications can be drawn from the results of this series of experiments? First, the similar patterns observed for uninflected nouns in sentence contexts (Experiment 6.1) and the same nouns in isolation (Experiment 6.2) confirm that the sentences were indeed constructed such that they did not introduce semantic constraints on the final nouns. Otherwise, different patterns should have been observed for nouns in contexts than for the same nouns in isolation. This suggests that in both cases, lexical information rather than sentential/semantic information biased the listeners' responses. Note also that the results obtained for nouns were quite consistent across experiments: the same pattern for nouns was also obtained in Experiment 4.1. The results in Experiment 5.1A, however, differed from all other results in that the strongest lexical effect was obtained in the medium RT range rather than in the fast RT range. This might have been due to the fact that in the fast RT range of Experiment 5.1A there were more responses in the verbal condition than in the nominal condition. This means that many fast responses were made to the verbal condition (due to an anticipatory strategy which will be discussed below). This might have pushed the lexical effect for nouns to the medium RT range. In other words, if there had not been an over-representation of verbal responses in the fast RT range in this experiment, the lexical effect for uninflected nouns might have been strongest in the fast RT range as well. Due to the high anticipation of the final word in the verbal condition, fast responses in that condition were unusually fast so that the fast responses in the nominal condition appear in a slower RT range relative to the verbal responses.

Second, inflected verbs presented in isolation produced reliable lexical effects (Experiment 6.2). However, this effect was weaker than that for uninflected nouns presented in isolation. This effect is in contrast to the hypothesis that the effect for inflected verbs might be stronger than for uninflected nouns due to the representation of morphological structure in verbs. If, for example, inflected verbs are represented in a morphologically decomposed fashion, with a separate representation of the morpheme '-t', there could have been a stronger bias towards [t] in inflected verbs as compared to uninflected nouns. This is not what was observed. What the results suggest instead is that the occurrence of inflected verbs in isolation is less common than the occurrence of isolated nouns. When inflected verbs are presented in isolation, there is no syntactic frame that requires or - linguistically speaking - licenses an inflectional morpheme. In principle, a Dutch verb presented in isolation could either be not inflected at all (i.e., 'ga' in 'ik ga', *I walk*) or end with either the inflectional marker '-n' (i.e., 'gaan' in ’wij/zij gaan’, *we/they walk*) or the inflectional marker ’-t’ (i.e., ’gaat’ in ’hij/zij gaat’, *he/she walks*). The lexical activation of ’-t’ in ’straat’ (*street*) might thus be somewhat stronger than the lexical activation of ’-t’ in ’gaat’. This might explain the more robust effect observed for nouns as compared to verbs. This furthermore rules out the hypothesis that there might be a special link between inflectional morphemes and decision units as had been suggested earlier. It is thus more likely that the lexical effects observed for both inflected verbs and uninflected nouns were mediated by full-form representations rather than by decomposed representations. It is also possible, however, that the stronger lexical effect for uninflected nouns is due to the factor frequency. Remember, that it was not possible to fully balance the surface frequencies of inflected verbs and uninflected nouns. Because nouns were on average more frequent than verbs, the lexical activation of ’-t’ as part of a noun might have been stronger than the lexical activation of ’-t’ as part of an inflected verb.

Third, the strong sentential/lexical effect that was observed for inflected verbs in sentence contexts in Experiment 5.1A (with many stimulus repetitions) vanished when more variability was introduced to the task in Experiment 6.1. This suggests that the results from Experiment 5.1A were driven by a *combination* of the high experimental predictability and the syntactic predictability of the final sound in the verbal sentence contexts. Experimental predictability was so strong that 10% of listeners’ fastest reactions to the [k] endpoint of the continuum (i.e., the unambiguous instantiation of [k]) were [t] responses in the verbal sentences. In some sense, the subjects’ expectation of the verb ’gaat’ was so high that their fastest reactions to an unambiguous [k] were still biased towards a [t] response. This interpretation is confirmed by the different result that was obtained in Experiment 6.1 where the final words were not so predictable. The strong effect for inflected verbs in appropriate sentence frames completely vanished in Experiment 6.1. However, in Experiment 6.2 there was a clear lexical effect for uninflected verbs in isolation that had not been observed for the same verbs in appropriate sentence contexts. Thus, verbs only showed lexical effects when they were presented in isolation. The lack of a “sentential/lexical” effect for inflected verbs in Experiment 6.1 might therefore be attributable to the processing of the preceding context. But how can such a “null” effect be attributed to the processing of the context?

Since the final sound in the verbal contexts was in a morphemic position (i.e., ’Vraag jij of Jan morgen vist?’, *Are you asking whether Jan fishes tomorrow?*), the subjects might have been able to separate this final sound from the rest of the sentence and perform the categorization task on the final sounds independently of the context. This decompositional strategy was only possible when inflected verbs were presented in sentence contexts. It was not possible for uninflected nouns presented either in sentence contexts or in isolation because in neither case was the final sound in a morphemic position. And this strategy was also not possible when inflected verbs were presented in isolation because there was no sentence context that allowed for a decompositional strategy. Such a dissociation of sentential and phonetic processing is in line with what Eimas et al. (1990) and Eimas and Nygaard (1992) found with the phoneme monitoring task. They also failed to observe lexical effects in sentence contexts and proposed therefore that listeners under some (experimental) circumstances might find it more economical to base two different decisions (the primary phonetic decision and the secondary lexical decision) on two different sources of information rather than consulting only one information-source for both decisions. The current data seems to suggest a similar mechanism: because the final phoneme in verbal sentence contexts was in a morpheme position, listeners were able to dissociate syntactic/sentential processing from phonetic processing. This interpretation is also in line with results reported by Miller et al. (1984) who showed that sentential effects on phonetic decision are not mandatory. Only when the listeners' attention to the sentential information was explicitly required to perform the task did Miller et al. (1984) observe lexical effects on phonetic decisions.

Note that these results (like those of Eimas & Nygaard, 1992) are direct consequences of the task listeners have to perform and might therefore not be directly informative about the natural processes of language comprehension. That is not to say, however, that the results are not meaningful with respect to the morphological structure of the mental lexicon. If the interpretation provided above is true, then the central decomposition of inflected verbs is obligatory in order to explain the different effects observed for uninflected nouns and inflected verbs. If there was no morphological information at that level, that is, if there were only full-form representations of inflected words, the decomposition of the final verbs would not have been possible, which in turn should have resulted in a similar effect for inflected verbs as for uninflected nouns (i.e., a lexical effect for inflected verbs in both sentence contexts and in isolation).

In sum, the results show that the responses to uninflected nouns presented both in sentential frames and in isolation as well as responses to verbs in isolation were influenced by *lexical* information, while *sentential* information only played a role when inflected verbs were presented in appropriate sentence contexts. This interpretation is however constrained in the sense that it is based on a null effect. A stronger conclusion would be possible if the differential effects for uninflected nouns and inflected verbs in contexts did not involve a null effect. It might be possible to avoid a decompositional strategy - as might have been used in Experiment 6.1 for inflected verbs - if the ambiguous sound in morpheme position was not presented at the very end of the experimental sentences. This could be achieved by presenting the sentential/syntactic information **after** the relevant sound rather than **before** this sound. Following syntactic information has an influence on the categorization of ambiguous sounds (van Alphen & McQueen, 2001). If this information came later than the to-be-categorized sound, there might not be enough time or sufficient syntactic constraints for listeners to be able to decompose inflected verbs that contained that sound.

Furthermore, it would be very interesting to compare sentential/semantic effects with sentential/syntactic effects: the nominal sentence contexts could be designed such that they strongly constrained the use of a certain noun (like the sentences in the studies by Miller et al. (1984) and Borsky et al., 1998). Would there be different patterns of results for ambiguous sounds that formed part of semantically predictable uninflected nouns than for the same ambiguous sounds as parts of syntactically predictable inflected verbs? If the dissociation interpretation given for inflected verbs in sentence contexts is indeed true, even a strong semantic bias in the nominal condition should not cause the lexical effect for uninflected nouns to vanish. Such a dissociation should only be possible when the to-be-categorized sound is separable from the context. Such a design might allow for a more direct comparison of uninflected nouns and inflected verbs in sentence contexts since the categorization of ambiguous sounds in both cases could be influenced by three processing levels: the phonemic level, the lexical level, and the sentential level.

Finally, would there be different effects for inflected verbs presented in sentence contexts when listeners were asked to perform a secondary task (as in the Miller et al., 1984, study) which explicitly required them to pay attention to the syntactic structure of the preceding (or following) sentence context? Clearly, many more studies can be envisaged that might allow for a better understanding of the relative contributions of semantic, syntactic and morphological information on phonetic decisions.

Summary and Conclusions

Chapter 7

In order to understand spoken language, we need to map the information in the continuous and highly variable speech stream onto stored lexical knowledge. This task is performed by a recognition system whose basic architecture consists of three processing levels: a *prelexical* level, at which perceptual units code information about the signal itself and which mediates between the incoming auditory signal and form representations of words located at the *lexical* level, and a *central* level where semantic, syntactic, pragmatic etc. features of words are represented. Word representations at the lexical level are also referred to as *access* representations that code either phonetic or orthographic features of words and which are assumed to be independent from *central* representations.

The current study focussed on two major issues in research of spoken language comprehension: segmentation and phonetic decision-making. Specifically, both series of experiments took a morphological perspective of these topics and tried to investigate the relationship between morphology and early processes of speech comprehension. Since morphology describes the link between form and meaning, it might have an important role to play at all three processing levels of the recognition system. A central issue in that respect is the form representations of morphologically complex words take: are these complex words stored in a full-form or decomposed fashion?

The role of morphology in the segmentation process

The research in the first part of the thesis (Chapters 2 and 3) sought a better understanding of the characteristics of one specific segmentation principle, the Possible Word Constraint (PWC; Norris et al., 1997). This constraint states that listeners segment the incoming speech such that no impossible words are left over. When the recognition system encounters the last two words (*Fall* [fʌl] and *streiten* [ʃtrɑɪtən]) in the German sentence *Man kann über diesen Fall streiten* (One can discuss about this case), words like *falsch* (wrong) and *reiten* (horse back riding) - among others - are temporarily activated but would lead to a segmentation that left the sound [t] unaccounted for. Since single consonants are not possible words in German (as is also the case in all other European languages) a parse that would leave a single consonant residue is penalized by the PWC. The PWC uses a variety of cues to the location of possible word boundaries such as metrical structure, phonotactic constraints, and acoustic information. While these cues are subject to language-specific variation, the PWC appears to be universal (McQueen et al., 2001; Norris et al., 2001; Cutler et al., submitted).

In its current version, the PWC states that any parse which results in a sequence of phonemes that does not contain a vowel should be penalized (e.g., *wach* ’awake’ in *schwach* ’weak’). For European languages in which single consonants can constitute morphemes (e.g., the ’-t’ in the Dutch verb *loopt*, ’she/he walks’) this constraint therefore has interesting consequences for the processing of inflected forms. Upon hearing an inflected form in the phrase *zij loopt* (she walks), the activation of the competing form *loop* will be attenuated by the PWC, since the ’-t’ between the *p* of *loop* and the end of the phrase is not a possible word. The operation of the PWC would thus guarantee successful recognition of the intended inflected word.

This would only be true, however, if there are full-form representations of morphologically complex words (e.g., of the inflected word *loopt*). If there were only decomposed representations of inflected words like *loopt*, successful recognition of *loopt* would have to be mediated via the stem *loop* and an access representation of the inflectional morpheme *-t*. In such a case a morphologically insensitive PWC would erroneously penalize the activation of the competing form *loop* because the constraint would not ’recognize’ the morphological status of the final *-t*. Recognition would not be impaired, however, if there were decomposed representations and the PWC was morphologically sensitive. Thus, if the PWC were found to be sensitive to morphological information, this would be in line with decompositional accounts of lexical access. Note, however, that a morphologically sensitive PWC would not rule out models with full-form representations of morphologically complex words. But there would then need to be a mechanism to resolve competition between, for example, *loopt* and *loop*. If single consonant morphemes are not treated differently during segmentation than morphologically meaningless consonants (i.e., if the PWC is insensitive to morphological information), on the other hand, this would be in favor of models with full-form access representations and would challenge models that require obligatory decomposition of complex words prior to lexical access.

In the first series of experiments (Chapters 2 and 3), the word-spotting task was used in order to address this issue. In Experiment 2.1A, monosyllabic Dutch words were embedded in three different kinds of nonsense strings. One was designed such that the detection of the embedded word was supposed to be relatively easy because the context that followed the embedded word was a syllable (e.g., *deur* ’door’ in *deurtach*; *syllabic* condition), that is, a possible word in Dutch. The second condition, on the other hand, was supposed to make the spotting of the embedded words difficult because the following contexts consisted of single consonants (e.g., *deur* in *deurp*; the *consonantal* condition). In the third condition, the consonants which followed the embedded words were inflectional morphemes of Dutch (i.e., either the verbal marker ’-t’ or the nominal marker ’-s’ [the *morphological* condition]; *deur* in *deurt* [the *stop* item set] or *duim* in *duims* [the *fricative* item set]). The logic was that the spotting of *deur* in *deurt* would be as easy as the spotting of *deur* in *deurtach* if the PWC was sensitive to morphological information. If this was not the case, the spotting of *deur* in *deurt* should be as difficult as in *deurp*.

Although Experiment 2.1A replicated earlier findings in English, for example, where word-spotting latencies were faster when the following context was a syllable than when the following context was a single consonant (Norris et al., 1997), no reliable differences were observed between the *morphological* condition (i.e., *deurt*) and either of the other two conditions. This was mainly attributed to two confounds: first, the small number of items that had been used (30 in total) had probably limited the experimental power, and second, the sequential probabilities at the offsets of nonsense strings in the *consonantal* condition of one item set (e.g., *duim* in *duimf*) were very low so that the task could have been solved on the basis of this information alone. Since listeners are sensitive to the likelihood of phoneme sequences in their languages (van der Lugt, 1999, 2001) a low frequent phoneme cluster like *mf* might make the location of the correct word boundary easy because the sequence *mf* hardly ever occurs at the offsets of Dutch words.

The two different item sets were therefore investigated in two separate experiments so that more experimental stimuli could be used (Experiment 3.1). The monosyllabic items that had been used in Experiment 2.1A were supplemented by bisyllabic nouns. Furthermore, sequential probabilities at the offsets of the nonsense strings in the *fricative* item set were now higher than in the first experiment. The aim of Experiment 3.1 was to obtain a clear result for the *morphological* condition. But what was observed instead was a reversal of the PWC effect, as established by Norris et al. (1997). That is, the spotting of words was faster in the *consonantal* condition than in the *syllabic* condition. Thus, the condition which was predicted to be the slowest condition produced the fastest RTs. Although the *morphological* condition clustered with the *consonantal* condition, an interpretation of that effect was not possible because the unexpected results in the *syllabic* condition indicated that there was a confound in the stimuli. A control experiment (3.2) investigated whether metrical differences between the conditions caused the reversed effect, but the same pattern of results emerged. Positive correlations of RTs with the lengths of the following contexts indicated that listeners did not answer as fast as possible but rather waited until the ends of the whole strings before they initiated their responses.

The following two Experiments (3.3 and 3.4) each introduced an extra factor in order to test possible explanations for the listeners' waiting strategy. The hypothesis was that listeners had been able to identify the embedded words before they encountered the following contexts that were supposed to influence word-spotting performance. This early identification might have been due to the early uniqueness points of the bisyllabic items, on the one hand, or the fixed embedding position of target words at the onsets of the nonsense strings, on the other hand. Thus, listeners could rely on the fact that the onsets of the embedded words always matched the onsets of the longer nonsense strings. Both factors made the word-spotting task so easy that listeners had in some sense the luxury of waiting until the ends of the whole strings before they gave their answers. Experiment 3.3 addressed this issue by including nonsense strings in the item lists which had targets embedded at their final position (e.g., *lepel* in *blepel* or *kulepel*). The idea was that the higher task demands (i.e., listeners could not assume that the targets were always in the same position) would delay the identification of the initially-embedded (i.e., the crucial) target words and hence would allow the following context to exert an influence on listeners' word-spotting performance. While items with finally-embedded targets produced a robust PWC effect (i.e., faster RTs for spotting *lepel* in *kulepel* than in *blepel*), there was again a reversed effect for targets embedded item-initially.

In Experiment 3.4, the lengths of the following contexts were balanced between the three conditions. These had varied in all previous experiments, since syllables are longer than single consonants. If the PWC effect was simply masked by the waiting strategy, that is, if the type of following context was effective but not visible in the results because of the waiting strategy, then equal context lengths should have helped to unmask the PWC effect. And indeed, the results obtained in the final word-spotting experiment produced an effect in the predicted direction: targets were spotted faster in the *syllabic* condition than in either of the other two conditions. That is, *probleem* (problem) was spotted faster in *probleem.dwaaf* than in *probleemt.daaf* or *probleemp.daaf*[1]. However, positive correlations of RTs with context lengths demonstrated once more that it was not the factor *context type* (i.e., *syllabic, consonantal, or morphological*) which de-

---

[1] Syllable boundaries based on phonotactic constraints are marked by a full stop.

termined the word-spotting latencies but rather the factor *context length*. The effect in Experiment 3.4 was simply in the predicted direction because now (in contrast to the other experiments) the mean context length in the *syllabic* (i.e., the fastest) condition was significantly shorter than in the other two conditions.

Unfortunately, the pattern of error rates also does not allow for a clear conclusion to be drawn about the role morphemes play in the segmentation process. Although only Experiment 3.3 showed a reversed effect for error rates comparable to the reversal observed for RTs (i.e., most errors in the *syllabic* condition, which was predicted to produce the lowest error rates), the error rates obtained in the *morphological* condition were not consistent across experiments.

In the General Discussion of Chapter 3 the current results were compared with results from other word-spotting studies which in part also obtained reversed PWC effects (e.g., McQueen & Cutler, in preparation). On the basis of these combined results, it was concluded that the appearance of a waiting-strategy depends on a fine balance of different design features such as the sequential probabilities of the target-bearing nonsense strings and the uniqueness points of the target words. Furthermore, the current data show that Dutch was not the ideal language to test these questions. Stimulus construction for word-spotting experiments in general is relatively constrained and the only Dutch words which met all these constraints carried the confounds which appeared to be responsible for the waiting strategy. It is thus possible that stimulus construction in another language, German for example, might be less constrained so that these confounds might be avoided. Yet another possibility might be to discourage the waiting strategy by making the following contexts so long (e.g., *deur* in *deursbugfum*) that the waiting strategy - presumably unconscious in the current study - would become so obvious that listeners might find it odd to wait until the end of each string.

Phonetic decision-making and morphological information

Phonetic decisions about ambiguous or distorted sounds are subject to influences from various sources of information. Best documented in the literature are lexical and sentential/semantic influences on listeners' phonetic categorizations of ambiguous phonemes while less work has been done on sentential/syntactic influences on phonemic decisions. To date, no study has investigated whether morphological information might interact with sentential/syntactic information during the process of phonemic decision-making. The second part of the thesis, therefore, addressed exactly this question: does morphological information contribute to listeners' responses when they are explicitly asked to identify an ambiguous sound as one of two phonemes? In seven phonetic categorization experiments, listeners were presented with coherent sentence contexts that either contained an inflected verb (e.g., *gaat* in *Vraag jij of Jan morgen gaat?* 'Are you asking whether Jan leaves tomorrow?') or an uninflected noun (e.g., *plaat* in *Zie nog wel eens een plaat?* 'Do you now and then see a record?'; Experiments 4.1, 5.1A, and 6.1). Both inflected verbs and uninflected nouns ended with the phoneme [t] which was in a morphemic position in the former but not in the latter case. Inflected verbs and uninflected nouns formed the word endpoints of word-nonword place-of-articulation continua which varied from a clear [t] (i.e., the word endpoint) to a clear [k] (i.e., the nonword endpoint). The core question of these experiments was whether morphological information would exert an influence on listeners' categorization performance when the ambiguous sounds were part of inflected verbs (i.e., in a morpheme position predictable from the sentence context) or when the same sounds were part of uninflected nouns (i.e., in a position not decomposable from the whole word). In order to be able to attribute potentially different results for inflected verbs and uninflected nouns to the processing of the preceding sentence contexts, the same words were also presented in isolation (Experiments 5.2, 5.3, and 6.2).

Furthermore, real verbs and nouns were substituted by pseudowords that had a real word embedded within them (e.g., *vlaa* 'custard' in *vlaat*: *Vraag jij of Jan morgen vlaat?* 'Are you asking whether Jan custards tomorrow?' and *Zie jij nog wel eens een vlaat?* 'Do you now and then see a custards?'). Would listeners' categorization performance be influenced by the fact that the decomposition of the pseudoword *vlaat* into its constituent parts *vlaa* and *'-t'* would only make sense in one of the two sentential surroundings (i.e., the verbal sentence frame)?

While a clear answer was possible to the last question about pseudowords, a more complex answer was needed to account for the different results that were obtained in Experiments 5.1A and 6.1 (i.e., inflected verbs and uninflected nouns presented in sentence contexts). From Experiment 5.1B (i.e., pseudowords in sentence frames) it appeared that pseudowords were not treated differently from nonwords. Categorization performance was the same whether the preceding context would have licensed the decomposition of the final pseudoword into a noun stem (i.e., *vlaa*) and the inflectional marker (i.e., *'-t'*) or not. Although it is apparent from advertisements (e.g., *"Cum laude gevlaait"*) that native speakers can analyse the morphological structure of novel forms like *gevlaait*, for example, and integrate these forms in a larger context, listeners' categorization decisions in the current study were not influenced by that ability.

Novel forms like *vlaat* certainly have the characteristic of not being well established forms in our mental lexicon. Therefore, the morphological analysis of such novel forms might not only take more time than the decompositional analysis of existing words but might moreover be more dependent on a coherent situational context that motivates the decomposition of these novel forms. Future categorization experiments with morphologically complex novel words might thus have to provide listeners with more situational information in order to test when and how they are decomposed during sentence comprehension.

As already mentioned above, the results for real verbs and nouns produced a more complex pattern. While the lexical/sentential effect for inflected verbs in sentence contexts was slightly stronger and faster than for uninflected nouns in sentence contexts when the **experimental** predictability was very high (Experiment 5.1A), the opposite result was observed when the experimental predictability was reduced (Experiment 6.1). More specifically, when both sentence frames and final words were more variable, uninflected nouns showed a clear lexical effect which was strongest in the listeners' fastest responses. No such effect was observed for inflected verbs. The strong effect that was observed for inflected verbs in sentence contexts in Experiment 5.1A was therefore mainly attributed to an anticipatory strategy of listeners due to a combination of the high proportion of stimulus repetitions, on the one hand, and the syntactic predictability of the final morpheme, on the other hand.

Interestingly, inflected verbs presented in isolation showed a different categorization pattern than the same verbs embedded in sentence frames (Experiment 6.2). When presented in isolation, inflected verbs showed a lexical effect which was confined to listeners' fastest responses. This lexical effect was slightly weaker for inflected verbs than for uninflected nouns, which, in contrast, showed an almost identical categorization pattern as the same nouns presented in sentences.

The correspondence of results for uninflected nouns presented both in coherent sentence structures and in isolation was taken as evidence that mainly *lexical* information determined listeners' categorization behaviour on nouns. This result was not surprising since the sentence contexts in the nominal condition did not put strong semantic constraints on the final nouns.

The different results obtained for inflected verbs presented in isolation as compared to the same verbs presented in sentence frames were interpreted as follows: the 'null' effect observed for inflected verbs presented in coherent sentences was attributed to a dissociation of sentential and phonetic processing which was only possible when the final sound was in a morpheme position. Such a dissociation of two different processing routines - under some (task-specific) circumstances - might be more economical for listeners than a unified processing approach when the phonetic decisions are based on higher order information (see also Eimas & Nygaard, 1992). In the case of uninflected nouns, however, this strategy was not available for the listeners because a separation (due to morphological decomposition) of the final sound from the final word was not possible. Note, however, that listeners' categorization performance on uninflected nouns was exclusively influenced by lexical information, since the sentential context was semantically neutral. A more direct test of the dissociation interpretation would be to investigate inflected verbs with final ambiguous sounds in sentence contexts that followed the critical word (see for example van Alphen & McQueen, 2001).

Morphological implications

The dissociation explanation for the results of Experiments 6.1 and 6.2 can of course only hold when one assumes that central decomposition of inflected verbs can take place. If this was not the case, that is, if the central level only contained full-form representations without coding morphological structure, a dissociation strategy should have been impossible for inflected verbs. It is thus only via decomposition that the phonetic decision to the final potential morpheme can be dissociated both from the processing of the stem and the processing of the sentence. Otherwise, the same pattern of results should have been observed for inflected verbs and uninflected nouns in coherent sentence structures.

When considering the same words in isolation, how does the weaker lexical effect for inflected verbs as compared to uninflected nouns fit into that interpretation? First, this weaker effect clearly speaks against a special link between morphemes and decision units which - if anything - should have resulted in a stronger effect for inflected verbs than for uninflected nouns. Instead, the weaker effect suggests that inflected verbs presented in isolation lack the sentential context which licenses the occurrence of an inflected form and are thus slightly odd word forms to consider in isolation. It is therefore possible that the lexical activation of [t] in an uninflected noun like *straat* might be stronger than the lexical activation of [t] in an inflected verb like *gaat*. But note, that this difference might also be due to the higher mean surface frequency of uninflected nouns as compared to that of inflected verbs. However, apart from the difference in robustness, both lexical effects showed the same time course since both effects were strongest in listeners' fastest decisions. It is therefore likely that inflected verbs in isolation were processed via full form representations at the access (i.e., the lexical) level.

These findings would thus imply that full form representations of morphologically complex words might be developed at the access level while there would be decomposition of these complex forms at the central level. Note, however, that this interpretation is partially based on a null effect (i.e., no lexical/sentential effect for inflected verbs in sentence contexts). Furthermore, the interpretation that concerns the central level is based on the processing of inflected verbs in sentence contexts while the interpretation that concerns the access level is based on the processing of verbs in isolation.

However, the current data add to what was concluded in the Introduction, where a brief summary of morphological studies was provided: the processing of morphologically complex forms is determined by various factors such as frequency and semantic transparency. The present results suggest in addition that processing of inflected forms via full-form or decomposed representations might in part depend on whether or not they are presented in coherent sentence structures.

Implications for the feedback debate and for models of speech recognition

Although the second part of the thesis was mainly focussed on morphological issues, the results make a small contribution to the feedback debate. Feedback, as it is currently incorporated in TRACE (McClelland & Elman, 1986), is the top-down flow of activation in the comprehension system. Due to the architecture of TRACE, activation which is fed back from the lexical level to the perceptual level (the phoneme level) accumulates over time. Thus, according to TRACE, any lexical or sentential effects on phonetic decisions should be stable over time or perhaps should tend to increase over time.

This prediction was not confirmed by the data at hand. The lexical/sentential effects observed in all experiments were strongest in listeners' fastest responses and started to die away once responses became slower. There was only one exception to that finding, namely the lexical effect observed for uninflected nouns presented in high predictable sentence frames (Experiment 5.1A), which was strongest in the medium RT range. This shift of the lexical effect for uninflected nouns might have been due, however, to the fact that a high proportion of fast responses in that experiment consisted of responses in the verbal condition. This proportional over-representation of verbal responses in the fast RT range in Experiment 5.1A might have pushed the lexical effect for uninflected nouns into the medium RT range. Moreover, although the lexical effect for uninflected nouns in highly predictable sentence frames was 'delayed', it did not build up over time (i.e., it was not present in the slowest responses) as TRACE predicts.

The time pattern of lexical (sentential) effects observed in the current categorization experiments is consistent with the predictions of the autonomous model Merge (Norris et al., 2000), which assumes information to flow strictly bottom-up in the system. The model assumes that information from the lexical or the sentential level can influence phonetic decisions - made at a dedicated decision level - only as long as lexical and/or sentential integration has not been completed. Since lexical and sentential integration in the current experiments coincided with the presentation of the final sound of each sentence (or word), lexical or sentential effects were predicted to be strongest in listeners' fastest responses. There was no following context which might have delayed the integration of lexical or sentential information. The time course of the effects in the present categorization experiments support the architecture of Merge in which all levels of information (i.e., phonemic, lexical, and sentential information) can feed information to the decision stage but can only be effective as long as the processing of the speech input has not been resolved (see also van Alphen & McQueen, 2001). Since the resolution of both lexical and sentential information was always possible shortly after the presentation of the final phoneme, higher-order effects were confined to the fastest responses.

Models that are concerned with spoken word recognition (e.g., TRACE and Shortlist) have to date not explicitly addressed morphological issues. These models thus make the simplifying assumption that all words are represented as full forms and that there is no morphological information at any level of the recognition system. As already mentioned in the Introduction, there is clear evidence that the processing of spoken language is sensitive to morphological information. Therefore models like TRACE or Shortlist should try to incorporate morphological information in an appropriate fashion. Unfortunately, the current data do not constrain such morphological implementations very strongly. Since the results of the word-spotting experiments in the first part of the thesis do not allow for any conclusions as to whether the PWC is morphologically sensitive or not, no adaptation of this constraint, as implemented in the Shortlist model, is necessary. If, however, future research might demonstrate that single consonant morphemes do have a special status in segmentation, then Shortlist might need to be modified. A simple modification would be to assume the decomposition of a complex form like *loopt* into its constituents parts so that this word would be recognized via the parallel activation of the stem *loop* and the inflectional morpheme *-t*. There would thus be no competition between the forms *loop* and *loopt*. A clear demonstration of the morphological insensitivity of the PWC, on the other hand, would show that the constraint in its current implemented version is correct.

The results of the categorization experiments in the second part of the thesis suggest that isolated inflected verbs might be processed just like uninflected nouns via full-form representations. Thus, so far, no additional features need to be represented at the lexical level of a model like Shortlist or TRACE. The results further suggest, however, that morphological decomposition might take place at the central level when inflected verbs are presented in an appropriate sentence contexts. If this is true, a model like Shortlist would need an additional level at which morphological information was coded.

In sum, a closer link between models of morphological processing - which have mainly been driven by research on visual word recognition - and models of spoken word recognition would provide a more complete account of speech comprehension. Thus, in future research, morphological issues should be more closely combined with the central issues of spoken word recognition: issues like segmentation, competition, and the high variability of the auditory signal.

References

Alfonso, P. J. (1981). *Context effects on the perception of place of articulation.* (Paper presented at the meeting of the Acoustical Society of America, Ottawa, Ontario, Canada.)

Alphen, P. van, & McQueen, J. M. (2001). The time-limited influence of sentential context on function word identification. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(5), 1057–1071.

Baayen, H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*, 94–117.

Baayen, H., McQueen, J. M., Dijkstra, T., & Schreuder, R. (2001). Dutch inflectional morphology in spoken- and written-word recognition. *Manuscript in preparation*.

Baayen, H., Piepenbrock, R., & Rijn, v. H. (1993). *The CELEX lexical database (CD-ROM)* (Tech. Rep.). Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Beckman, M. E., & Edwards, J. (1990). Lengthenings and shortenings and the nature of prosodic constituency. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 152–178). Cambridge: Cambridge University Press.

Bertram, R., Laine, M., & Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphologically rich language. *Journal of Psycholinguistic Research*, *28*, 213–226.

Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 489–511.

Borsky, S., Tuller, B., & Shapiro, L. P. (1998). ”How to milk a coat”: The effects of semantic and acoustic information on phoneme categorization. *Journal of the Acoustical Society of America*, *103*, 2670–2676.

Bradley, D. C., Sánchez-Casas, R. M., & García-Albea, J. E. (1993). The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes*, *8*, 197–233.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production* (Vol. II, pp. 257–332). London: Academic Press.

Cairns, P., Shillcock, R., Charter, N., & Levy, J. (1995). Bottom-up connectionst modelling of speech. In J. P. Levy, D. Bairaktaris, & J. A. B. P. Cairns (Eds.), *Connectionst models of memory and language* (pp. 289–310). London: UCL Press.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*, 291-332.

Clahsen, H. (1996). The perception of participles in the german mental lexicon: Evidence for the dual-mechanism model. In G. Booij & J. van Merle (Eds.), *Yearbook of morphology* (pp. 117–134). Dordrecht: Kluwer.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of german inflection. *Behavioral and Brain Sciences*, *22*, 991–1060.

Clahsen, H., Eisenbeiss, S., & Sonnenstuhl-Henning, I. (1997). Morphological structure and the processing of inflected words. *Theoretical Linguistics*, *23*, 201–249.

Connine, C. M. (1987). Constraints on interactive processes in auditory word recognition: The role of sentence context. *Journal of Memory and Language*, *26*, 527–538.

Connine, C. M. (1990). Effects of sentence context and lexical knowledge in speech processing. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 281–294). Cambridge, MA: MIT Press.

Connine, C. M., Blasko, D. G., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints. *Journal of Memory and Language*, *30*, 234–250.

Connine, C. M., & Clifton, C. J. (1987). Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(2), 291–299.

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 1084–1096.

Connine, C. M., & Titone, D. (1997). Phoneme monitoring. In F. Grosjean & U. H. Frauenfelder (Eds.), *A guide to spoken word recognition paradigms* (pp. 635–647). Sussex, UK: Psychology Press.

Connine, C. M., Titone, D., Deelman, T., & Blasko, D. (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language*, *37*, 463–480.

Connine, C. M., Titone, D., & Wang, J. (1993). Auditory word recognition: Extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(1), 81–94.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, *31*, 218–236.

Cutler, A., Demuth, C., & McQueen, J. M. (submitted). Universality versus language-specificity in listening to running speech.

Cutler, A., Mehler, J., Norris, D., & Seguí, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*, 385–400.

Cutler, A., Mehler, J., Norris, D., & Seguí, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, *19*, 141–177.

Cutler, A., & Norris, D. (1979). Monitoring sentence comprehension. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett.* Erlbaum.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 113–121.

Drews, E. (1997). Morphological priming. In F. Grosjean & U. H. Frauenfelder (Eds.), *A guide to spoken word recognition paradigms* (pp. 629–634). East Sussex: Psychology Press.

Eimas, P. D., & Corbit, J. D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, *4*, 99–109.

Eimas, P. D., Hornstein, S. B. M., & Payton, P. (1990). Attention and the role of dual codes in phoneme monitoring. *Journal of Memory and Language*, *29*, 160–180.

Eimas, P. D., & Nygaard, L. C. (1992). Contextual coherence and attention in phoneme monitoring. *Journal of Memory and Language*, *31*, 375–395.

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, *27*, 143–165.

Fox, R. A. (1984). Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(4), 526–540.

Frauenfelder, U. H., Seguí, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 77–91.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 110–125.

Gaskell, G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and cognitive processes*, *12*, 613–656.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251–79.

Gow, D., & Gordon, P. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 344–359.

Isenberg, D., Walker, E. C. T., & Ryder, J. M. (1980). A top-down effect in the identification of function words. *Paper presented at the Meeting of the Acoustical Society of America, Los Angeles, California*.

Jarvella, R., & Meijers, G. (1983). Recognizing morphemes in spoken words: Some evidence for a stem-organized mental lexicon. In G. D'Arcais & R. Jarvella (Eds.), *The process of language understanding* (pp. 81–112). New York: Wiley.

Jong, N. de, Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processing*, *15*, 329–365.

Karlsson, F., & Koskenniemi, K. (1985). A process model of morphology and lexicon. *Folia Linguistica*, *29*, 207–231.

Lehiste, I. (1960). An acoustic-phonetic study of internal open juncture. *Phonetica*, *5*, 1–54.

Luce, P., Pisoni, D., & Goldinger, S. (1990). Similarity neighborhoods of spoken words. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 122–147). Cambridge, MA: MIT Press.

Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation and PARSYN. *Perception and Psychophysics*, *62*, 615–625.

Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes*, *16*, in press.

Lugt, A. v. d. (1999). *From speech to words.* Unpublished doctoral dissertation, Katholieke Universiteit Nijmegen, Nijmegen.

Lugt, A. v. d. (2001). The use of sequential probabilities in the segmentation of speech. *Perception and Psychophysics*, *63*(5), 811–823.

Mann, V. A., & Repp, B. H. (1981). Influence of preceding fricative on stop consonant perception. *Journal of the Acoustical Society of America*, *69*, 548–558.

Marslen-Wilson, W., Komisarjevsky Tyler, L., Waksler, R., & Older, L. (1994). Morphology and meaning in the english mental lexicon. *Psychological Review*, *101*, 3–33.

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, *101*(4), 653–675.

Marslen-Wilson, W., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.

McQueen, J. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, *39*, 21–46.

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

McQueen, J., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 621–638.

McQueen, J., Otake, T., & Cutler, A. (2001). Rhythmic cues and possible-word constraints in Japanese speech segmentation. *Journal of Memory and Language*, *45*, 103–132.

McQueen, J. M. (1991). The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 433–443.

McQueen, J. M. (1996). Word spotting. *Language and Cognitive Processes*, *11*, 695–699.

McQueen, J. M., & Cutler, A. (1998). Spotting (different types of) words in (different types of) context. *Proceedings of the Sixth International Conference on Spoken Language Processing*, *6*, 2791–2794.

McQueen, J. M., & Cutler, A. (in preparation). The PWC and phonotactics.

McQueen, J. M., Norris, D., & Cutler, A. (1999). Lexical influence in phonetic decision making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(5), 1363–1389.

Mehler, J., Dommergues, J.-Y., Frauenfelder, U., & Seguí, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal*

Behavior, *20*, 298–305.

Miller, J. L., & Dexter, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 369–378.

Miller, J. L., Green, K., & Schermer, T. (1984). A distinction between the effects of sentential speaking rate and semantic congruity on word identification. *Perception and Psychophysics*, *36*, 329–337.

Newman, R. S., Sawusch, J. R., & Luce, P. A. (1997). Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 873–889.

Niemi, J., Laine, M., & Tuominen, J. (1994). Cognitive morphology in Finnish: Foundations of a new model. *Language and Cognitive Processes*, *9*, 423–446.

Norris, D. (1982). Autonomous processes in comprehension: A reply to Marslen-wilson and Tyler. *Cognition*, *11*, 97–101.

Norris, D. (1993). Bottom-up connectionist models of ”interaction”. In G. T. M. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The second Sperlonga meeting.* Erlbaum.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234.

Norris, D., McQueen, J., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1209–1228.

Norris, D., McQueen, J., Cutler, A., Butterfield, S., & Kearns, R. (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes*, *16*, in press.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging phonetic and lexical information in phonetic decision-making. *Behavioral and Brain Sciences*, *23*, 299–325.

Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, *34*, 191–243.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, *32*, 358–378.

Pitt, M. A. (1995). The locus of the lexical shift in phoneme identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1037–1052.

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, *39*, 347–370.

Pitt, M. A., & Samuel, A. (1995). Lexical and sublexical feedback in auditory word recognition. *Cognitive Psychology*, *29*, 149–188.

Pitt, M. A., & Samuel, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 699–725.

Quené, H. (1992). Durational cues for word segmentation in Dutch. *Journal of Phonetics*, *20*, 331–350.

Repp, B. H. (1981). Perceptual equivalence of two kinds of ambiguous speech stimuli. *Bulletin of the Psychonomic Society*, *18*, 12–14.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, *110*, 474–494.

Samuel, A. G. (1997). Lexical activation produces potent phonemic percepts. *Cognitive Psychology*, *32*, 97–127.

Samuel, A. G. (2001). Some empirical tests of Merge's architecture. *Language and Cognitive Processes*, *16*, in press.

Sawusch, J. R., & Jusczyk, P. (1981). Adaptation and contrast in the perception of voicing. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 408–421.

Schreuder, R., & Baayen, H. (1995). Modeling morpholgical processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Erlbaum.

Schriefers, H., Zwitserlood, P., & Roelofs, A. (1991). The identification of morphologically complex spoken words: Continuous processing or decomposition. *Journal of Memory and Language*, *30*, 26–47.

Sebastian-Gallés, N., Dupoux, E., Seguí, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, *31*, 18–32.

Seguí, J., Frauenfelder, U., & Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology*, *72*, 471–477.

Smits, R. (2000). Temporal distribution of information for human consonant regognition in VCV utterances. *Journal of Phonetics*, *27*, 111–135.

Stevenson, D. C. (1979). *Categorical perception and selective adaptation phenomena in speech.* Unpublished doctoral dissertation, University of Alberta, Edmonton, Alberta, Canada.

Suomi, K., McQueen, J., & Cutler, A. (1997). Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language*, *36*, 422–444.

Tabossi, P. (1997). Cross-modal semantic priming. In F. Grosjean & U. H. Frauenfelder (Eds.), *A guide to spoken word recognition paradigms* (pp. 569–576). East Sussex: Psychology Press.

Taft, M. (2001). *Does the lack of a base frequency effect necessarily imply direct whole word access?* (Paper presented at the Workshop 'Perspectives on Morphological Processing II'. Nijmegen, The Netherlands.)

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behaviour*, *14*, 630–647.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in spoken word recognition. *Psychological Science*, *9*, 325–329.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognitiion. *Journal of Memory and Language*, *40*, 374–408.

Vroomen, J., & Gelder, B. de. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 98–108.

Vroomen, J., Tuomainen, J., & Gelder, B. de. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language*, *38*, 133–149.

Vroomen, J., Zon, M. van, & Gelder, B. de. (1996). Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. *Memory and Cognition*, *24*, 744–755.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*, 392–393.

Weber, A. (2001). *Language specific listening: The case of phonetic sequences.* Unpublished doctoral dissertation, Katholieke Universiteit Nijmegen, Nijmegen.

Wurm, L., & Ross, S. E. (2001). Conditional root uniqueness points: Psychological validity and perceptual consequences. *Journal of Memory and Language*, *45*, 39–57.

Wurm, L. H., & Samuel, A. G. (1997). Lexical inhibition and attentional allocation during speech perception: Evidence from phoneme monitoring. *Journal of Memory and Language*, *36*, 165–187.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word recognition. *Cognition*, *32*, 25–64.

Zwitserlood, P. (1997). Form priming. In F. Grosjean & U. H. Frauenfelder (Eds.), *A guide to spoken word recognition paradigms* (pp. 589–595). East Sussex: Psychology Press.

Zwitserlood, P. (2001). *Comparing priming effects within and across morphological classes.* (Paper presented at the Workshop ’Perspectives on Morphological Processing II’. Nijmegen, The Netherlands.)

Materials

Appendix A

Materials for Experiment 2.1A

Experimental items by type of context. The target words are shown in upper case and the items are given in orthographic transcription.

Stop item set

morphological context (-t)	consonantal context	syllabic context	target translation
BOELt	BOELp	BOELpuis	whole lot
DEURt	DEURp	DEURtach	door
GANGt	GANGk	GANGkor	corridor
GEULt	GEULp	GEULpif	channel
HAMt	HAMp	HAMpijk	ham
NORt	NORp	NORpuul	prison
NULt	NULk	NULtoof	zero
PEULt	PEULk	PEULkaam	pod
RUMt	RUMp	RUMpien	rum
TANGt	TANGk	TANGkuuf	tongs
TORt	TORk	TORpig	beetle
WANGt	WANGk	WANGteul	cheek
ZANGt	ZANGk	ZANGtijn	song/singing
GONGt	GONGk	GONGkel	gong
BILt	BILp	BILtaag	buttock

Materials for Experiment 2.1A (continued)

Fricative item set

morphological context (-s)	consonantal context	syllabic context	target translation
DUIMs	DUIMf	DUIMfoel	thumb
GEURs	GEURf	GEURfum	odour
GIERs	GIERf	GIERsor	vulture
GYMs	GYMf	GYMseel	gym
KEELs	KEELf	KEELsuuf	throat
KIEMs	KIEMf	KIEMfag	germ
KIERs	KIERf	KIERfuim	chink
MOLs	MOLf	MOLsief	mole
ROLs	ROLf	ROLfiet	roll
VIERs	VIERf	VIERsuik	four
ZEURs	ZEURf	ZEURsif	bore
SCHUIMs	SCHUIMf	SCHUIMsuuk	foam
KLIMs	KLIMf	KLIMfas	climb
SCHOLs	SCHOLf	SCHOLfem	school
VORMs	VORMf	VORMfiek	form

Materials for Experiment 2.1A (continued)

Extra target-bearing items

syllabic context	target translation
BALfoeg	ball
BELseng	bell
DALpuuf	valley
HALteup	hall
HEERfog	army
BARkuuk	bar
JAARpeek	year
KUILsuug	club
TAALpeeg	language
MUURtaam	wall

Materials for Experiment 3.1

Experimental items by type of context. The target words are shown in upper case and the items are given in orthographic transcription.

Stop item set: bisyllabic targets

morphological context (-t)	consonantal context	syllabic context	target translation
PAROOLt	PAROOLk	PAROOLkiem	slogan
KOTTUUMt	KOSTUUMp	KOSTUUMtif	costume
MILIEUt	MILIEUk	MILIEUtuuf	environment
TABOEt	TABOEp	TABOEtijk	taboo
MOBIELt	MOBIELk	MOBIELtaaf	mobile
FLUWEELt	FLUWEELk	FLUWEELtieg	velvet
PROBLEEMt	PROBLEEMp	PROBLEEMtis	problem
VIOOLt	VIOOLk	VIOOLtem	violin
PARFUMt	PARFUMp	PARFUMpoek	perfume
KOPIEt	KOPIEk	KOPIEkel	copy
REPTIELt	REPTIELk	REPTIELket	reptile
VENTIELt	VENTIELk	VENTIELkuus	valve
SYSTEEMt	SYSTEEMp	SYSTEEMpon	system
TENEURt	TENEURk	TENEURtoop	tenor
RIOOLt	RIOOLk	RIOOLtuik	banner
PISTOOLt	PISTOOLk	PISTOOLtus	gun
NIVEAUt	NIVEAUp	NIVEAUpuul	level
LAURIERt	LAURIERk	LAURIERtech	laurel
TEXTIELt	TEXTIELk	TEXTIELkem	textile
JUWEELt	JUWEELk	JUWEELkog	jewel
MODELt	MODELk	MODELtung	model

Materials for Experiment 3.1 (continued)

Stop item set: monosyllabic targets

morphological context (-t)	consonantal context	syllabic context	target translation
DEURt	DEURp	DEURtach	door
NULt	NULk	NULtoof	zero
GANGt	GANGk	GANGkor	corridor
GEULt	GEULp	GEULpif	channel
BILt	BILp	BILtaag	buttock
ZANGt	ZANGk	ZANGtijn	song/singing
TORt	TORk	TORkig	beetle
PEULt	PEULk	PEULkaam	pod
WANGt	WANGk	WANGteul	cheek

Materials for Experiment 3.1 (continued)

Fricative item set: bisyllabic targets

morphological context (-s)	consonantal context	syllabic context	target translation
PAROOLs	PAROOLf	PAROOLfiem	slogan
FLUWEELs	FLUWEELf	FLUWEELfim	velvet
KASTEELs	KASTEELf	KASTEELfoes	castle
FORELs	FORELf	FORELsuil	trout
VENTIELs	VENTIELf	VENTIELfog	valve
GARNAALs	GARNAALf	GARNAALsuim	shrimp
SIGAARs	SIGAARk	SIGAARkief	cigar
REPTIELs	REPTIELf	REPTIELfin	reptile
VIOOLs	VIOOLf	VIOOLsem	violin
JUWEELs	JUWEELf	JUWEELfon	jewel
MAKREELs	MAKREELf	MAKREELsuin	mackerel
TEXTIELs	TEXTIELf	TEXTIELsan	textile
SYSTEEMs	SYSTEEMp	SYSTEEMpif	system
GITAARs	GITAARk	GITAARkuum	guitar
MOBIELs	MOBIELf	MOBIELfaaf	mobile
LAURIERs	LAURIERk	LAURIERkuul	laurel
RIOOLs	RIOOLf	RIOOLsim	banner
KANEELs	KANEELf	KANEELsiem	cinnamon
PISTOOLs	PISTOOLf	PISTOOLsam	gun
MODELs	MODELf	MODELsung	model
PROBLEEMs	PROBLEEMp	PROBLEEMpuif	problem
SPIRAALs	SPIRAALf	SPIRAALfich	spiral
KANTOORs	KANTOORk	KANTOORkig	office
TENEURs	TENEURk	TENEURkem	tenor

Materials for Experiment 3.1 (continued)

Fricative item set: monosyllabic targets

morphological context (-s)	consonantal context	syllabic context	target translation
DEURs	DEURk	DEURkach	door
GANGs	GANGk	GANGkor	corridor
ZANGs	ZANGk	ZANGkum	song/singing
GEULs	GEULf	GEULfip	channel
SCHOLs	SCHOLf	SCHOLfem	school
KIERs	KIERf	KIERfuim	chink
NULs	NULf	NULsoof	zero
KEELs	KEELf	KEELsuuf	throat
ZEURs	ZEURf	ZEURsif	bore
WANGs	WANGk	WANGkeul	cheek
SCHUIMs	SCHUIMp	SCHUIMsuuk	foam
PEULs	PEULf	PEULfaam	pod
MOLs	MOLf	MOLsief	mole
GIERs	GIERk	GIERsor	vulture
HAMs	HAMp	HAMpijk	ham
VORMs	VORMp	VORMsiek	form
GEURs	GEURk	GEURkum	odour
TANGs	TANGk	TANGkuuf	tong
DUIMs	DUIMp	DUIMpoel	thumb
ROLs	ROLf	ROLfiet	roll
BILs	BILf	BILsaag	buttock
VIERs	VIERf	VIERsuik	four
GONGs	GONGk	GONGsel	gong
KLIMs	KLIMp	KLIMpas	climb

Materials for Experiment 3.1 (continued)

Extra target-bearing items

syllabic context	target translation
BALfoeg	ball
BELseng	bell
DALpuuf	valley
HALteup	hall
HEERfog	army
BARkuuk	bar
JAARpeek	year
KUILsuug	club
TAALpeeg	language
MUURtaam	wall

Materials for Experiment 3.3

Experimental items used in both stop and fricative item sets by type of context. The target words are shown in upper case and the items are given in orthographic transcription.

Finally embedded targets

consonantal (C) context	reduced syllable (Cə) context	full syllable (CV) context	target translation
sLEZER	keLEZER	geeLEZER	reader
sLEUNEN	peLEUNEN	saLEUNEN	to lean
bLOPER	keLOPER	zoeLOPER	runner
sMELDEN	keMELDEN	huMELDEN	to mention
fREGIE	neREGIE	koREGIE	direction
sTEUGEL	feTEUGEL	fuTEUGEL	rein
sTITEL	keTITEL	seuTITEL	title
kNOEMEN	keNOEMEN	fuNOEMEN	to name
fREVUE	feREVUE	saREVUE	revue
kLEGER	keLEGER	neuLEGER	army
sMORREN	peMORREN	huiMORREN	to grumble
kWONING	feWONING	lieWONING	dwelling
dRICHEL	seRICHEL	baRICHEL	ledge
sMONNIK	seMONNIK	heuMONNIK	monk
sPUZZEL	fePUZZEL	geePUZZEL	puzzle
tREPLIEK	keREPLIEK	daREPLIEK	retort
bLEPEL	seLEPEL	kuLEPEL	spoon
sNUCHTER	feNUCHTER	duNUCHTER	sober
tROKEN	leROKEN	voeROKEN	to smoke
sMOLEN	keMOLEN	zaMOLEN	mill
sLIEGEN	feLIEGEN	baLIEGEN	to lie
fRILLEN	neRILLEN	goRILLEN	to shiver
bREKKEN	leREKKEN	noeREKKEN	to prolong
peTONEN	moTONEN	sTONEN	to show

Finally embedded targets (continued).

consonantal (C) context	reduced syllable (Cə) context	full syllable (CV) context	target translation
dWONEN	keWONEN	duWONEN	to live
sPUBER	sePUBER	goePUBER	adolescent
sMEUBEL	neMEUBEL	vieMEUBEL	furniture
sPEPER	kePEPER	foPEPER	pepper
gLITER	leLITER	goLITER	liter
sTAFEL	keTAFEL	foeTAFEL	black board
bRUBBER	feRUBBER	keeRUBBER	rubber
zWERPEN	keWERPEN	foeWERPEN	to throw
tWUIVEN	leWUIVEN	veeWUIVEN	to wave
fLEUGEN	seLEUGEN	toLEUGEN	lie
dREFREIN	seREFREIN	soeREFREIN	chorus
pSELECT	leSELECT	feuSELECT	to select
kRIDDER	keRIDDER	vijRIDDER	knight
sNUMMER	peNUMMER	voNUMMER	number
vRIMPEL	seRIMPEL	zaRIMPEL	wrinkle
sTUNNEL	seTUNNEL	soeTUNNEL	tunnel
gLOKKEN	keLOKKEN	fijLOKKEN	to lure
fJEUKEN	neJEUKEN	beeJEUKEN	to itch
bRECEPT	keRECEPT	kaRECEPT	recipe
gLACHEN	feLACHEN	deuLACHEN	to laugh
bLEZEN	keLEZEN	goLEZEN	to read
sMODDER	keMODDER	buMODDER	mud
sTREUREN	leTREUREN	saTREUREN	to mourn
kRECORD	peRECORD	fijRECORD	record

Materials for Experiment 3.4

Experimental items by type of context. The target words are shown in upper case and the items are given in orthographic transcription.

Stop item set: bisyllabic targets

morphological context (-t)	consonantal context	syllabic context	target translation
PAROOLtbiem	PAROOLkbiem	PAROOLbliem	slogan
KOTTUUMtdif	KOSTUUMpdif	KOSTUUMdwif	costume
MILIEUtbuuf	MILIEUkbuuf	MILIEUbluuf	environment
TABOEtdeek	TABOEpdeek	TABOEdreek	taboo
MOBIELtbaaf	MOBIELkbaaf	MOBIELblaaf	mobile
FLUWEELtbieg	FLUWEELkbieg	FLUWEELbrieg	velvet
PROBLEEMtdaaf	PROBLEEMpdaaf	PROBLEEMdwaaf	problem
VIOOLtbem	VIOOLkbem	VIOOLblem	violin
PARFUMtdem	PARFUMpdem	PARFUMdwem	perfume
KOPIEtbech	KOPIEkbech	KOPIEblech	copy
REPTIELtbeum	REPTIELkbeum	REPTIELbreum	reptile
VENTIELtbuus	VENTIELkbuus	VENTIELbruus	valve
SYSTEEMtdin	SYSTEEMpdin	SYSTEEMdrin	system
TENEURtboop	TENEURkboop	TENEURbroop	tenor
RIOOLtbuuk	RIOOLkbuuk	RIOOLbluuk	banner
PISTOOLtdes	PISTOOLkdes	PISTOOLdres	gun
NIVEAUtbuul	NIVEAUpbuul	NIVEAUbluul	level
LAURIERtdech	LAURIERkdech	LAURIERdwech	laurel
TEXTIELtbim	TEXTIELkbim	TEXTIELbrim	textile
JUWEELtbor	JUWEELkbor	JUWEELblor	jewel
MODELtbung	MODELkbung	MODELbrung	model

Stop item set: monosyllabic targets

morphological context (-t)	consonantal context	syllabic context	target translation
DEURtboch	DEURpboch	DEURbroch	door
NULtbool	NULkbool	NULbool	zero
GANGtbor	GANGkbor	GANGblor	corridor
GEULtbif	GEULpbif	GEULblif	channel
BILtbeem	BILpbeem	BILbreem	buttock
ZANGtdoon	ZANGkdoon	ZANGdroon	song/singing
TORtbis	TORkbis	TORbris	beetle
PEULtbuin	PEULkbuin	PEULbluin	pod
WANGtbeif	WANGkbeif	WANGbreif	cheek

Materials for Experiment 3.4 (continued)

Fricative item set: bisyllabic targets

morphological context (-s)	consonantal context	syllabic context	target translation
PAROOLsdul	PAROOLfdul	PAROOLdwul	slogan
FLUWEELsbim	FLUWEELfbim	FLUWEELbrim	velvet
KASTEELsdees	KASTEELfdees	KASTEELdrees	castle
FORELsdur	FORELfdur	FORELdrur	trout
VENTIELsbog	VENTIELfbog	VENTIELbrog	valve
GARNAALsdim	GARNAALfdim	GARNAALdrim	shrimp
SIGAARsbuif	SIGAARkbuif	SIGAARbluif	cigar
REPTIELsdin	REPTIELfdin	REPTIELdrin	reptile
VIOOLsdem	VIOOLfdem	VIOOLdwem	violin
JUWEELsdin	JUWEELfdin	JUWEELdrin	jewel
MAKREELsbuin	MAKREELfbuin	MAKREELbluin	mackerel
TEXTIELsduun	TEXTIELfduun	TEXTIELdruun	textile
SYSTEEMsdif	SYSTEEMpdif	SYSTEEMdwif	system
GITAARsbuum	GITAARkbuum	GITAARbruum	guitar
MOBIELsbaaf	MOBIELfbaaf	MOBIELblaaf	mobile
LAURIERsduul	LAURIERkduul	LAURIERdruul	laurel
RIOOLsbim	RIOOLfbim	RIOOLblim	banner
KANEELsbis	KANEELfbis	KANEELbris	cinnamon
PISTOOLsbam	PISTOOLfbam	PISTOOLblam	gun
MODELsbung	MODELfbung	MODELbrung	model
PROBLEEMsdaaf	PROBLEEMpdaaf	PROBLEEMdwaaf	problem
SPIRAALsbif	SPIRAALfbif	SPIRAALblif	spiral
KANTOORsbog	KANTOORkbog	KANTOORbrog	office
TENEURsbem	TENEURkbem	TENEURblem	tenor

Materials for Experiment 3.4 (continued)

Fricative item set: monosyllabic targets

morphological context (-s)	consonantal context	syllabic context	target translation
DEURsbug	DEURkbug	DEURblug	door
GANGsbor	GANGkbor	GANGbror	corridor
ZANGsdon	ZANGkdon	ZANGdwon	song/singing
GEULsdef	GEULdeff	GEULdwef	channel
SCHOLsdul	SCHOLfdul	SCHOLdwul	school
KIERsbuin	KIERfbuin	KIERbluin	chink
NULsboof	NULfboof	NULdwoof	zero
KEELsduuf	KEELfduuf	KEELdruuf	throat
ZEURsdif	ZEURfdif	ZEURdwif	bore
WANGsdeul	WANGkdeul	WANGdreul	cheek
SCHUIMsduuk	SCHUIMpduuk	SCHUIMdwuuk	foam
PEULsbam	PEULfbam	PEULbram	pod
MOLsbuf	MOLfbuf	MOLbruf	mole
GIERsbor	GIERkbor	GIERbror	vulture
HAMsdup	HAMpdup	HAMdwup	ham
VORMsdieg	VORMpdieg	VORMdrieg	form
GEURsbuif	GEURkbuif	GEURbluif	odour
TANGsbuuf	TANGkbuuf	TANGbruuf	tong
DUIMsbeum	DUIMpbeum	DUIMbreum	thumb
ROLsbif	ROLfbif	ROLblif	roll
BILsbeeg	BILfbeeg	BILbleeg	buttock
VIERsdeek	VIERfdeek	VIERdreek	four
GONGsdul	GONGkdul	GONGdwul	gong
KLIMsdes	KLIMpdes	KLIMdres	climb

Materials for Experiment 4.1

The experimental stimuli are listed per item set and the translations are given in italics.

Monosyllabic item set

lexical condition	verbal context (VP)
real word-nonword	de tante **gaat-gaak** *the aunt leaves*
pseudoword-nonword	de tante **vlaat-vlaak** *the aunt custards*
nonword-nonword	de tante **klaat-klaak** *the aunt klaat*

lexical condition	nominal context (NP)
real word-nonword	een brede **straat-straak** *a broad street*
pseudoword-nonword	een brede **vlaat-vlaak** *a broad custards*
nonword-nonword	een brede **klaat-klaak** *a broad klaat*

Materials for Experiment 4.1 (continued)

Bisyllabic item set

lexical condition	verbal context (VP)
real word-nonword	de moeder **geniet-geniek**
	the mother enjoys
pseudoword-nonword	de moeder **magiet-magiek**
	the mother magics
nonword-nonword	de moeder **keliet-keliek**
	the mother keliet

lexical condition	nominal context (NP)
real word-nonword	een stoere **bandiet-bandiek**
	a cool bandit
pseudoword-nonword	een stoere **magiet-magiek**
	a cool magics
nonword-nonword	een stoere **keliet-keliek**
	a cool keliet

Materials for Experiments 5.1 - 5.3

The experimental stimuli are listed per item set and translations are given in italics.

Monosyllabic item set

lexical condition	verbal context (VP)
real word-nonword	Vraag jij of Jan morgen **gaat-gaak**?
	Are you asking whether Jan leaves tomorrow?
pseudoword-nonword	Vraag jij of Jan morgen **vlaat-vlaak**
	Are you asking whether Jan custards tomorrow?
nonword-nonword	Vraag jij of Jan morgen **snaat-snaak**?
	Are you asking whether Jan snaat tomorrow?

lexical condition	nominal context
real word-nonword	Zie jij nog wel eens een **plaat-plaak**?
	Do you now and then see a record?
pseudoword-nonword	Zie jij nog wel eens een **vlaat-vlaak**
	Do you now and then see a custards?
nonword-nonword	Zie jij nog wel eens een **snaat-snaak**?
	Do you now and then see a snaat?

Materials for Experiments 5.1 - 5.3 (continued)

Bisyllabic item set

lexical condition	verbal context (VP)
real word-nonword	Lach jij als John haar **bespiedt-bespiek**? *Are you laughing when John spies on her*?
pseudoword-nonword	Lach jij als John haar **magiet-magiek** *Are you laughing when John magics her*?
nonword-nonword	Lach jij als John haar **meliet-meliek**? *Are you laughing when John meliet her*?

lexical condition	nominal context
real word-nonword	Wil jij een vloer van **graniet-graniek**? *Do you want a floor made from granite*?
pseudoword-nonword	Wil jij een vloer van **magiet-magiek** *Do you want a floor made from magics*?
nonword-nonword	Wil jij een vloer van **meliet-meliek**? *Do you want a floor made from meliet*?

Materials for Experiments 5.1 - 5.3 (continued)

Filler sentences

lexical condition	verbal context (VP)
real word-nonword	Vraag jij of Jan morgen **belt-belk**?
	Are you asking whether Jan rings tomorrow?
nonword-nonword	Vraag jij of Jan morgen **krelt-krelk**?
	Are you asking whether Jan krelt tomorrow?
real word-nonword	Lach jij als John haar **vergeet-vergeek**?
	Are you laughing when John forgets her?
nonword-nonword	Lach jij als John haar **bevleet-bevleek**?
	Are you laughing when John bevleet her?
real word-nonword	Huil jij als Daan met hem **spot-spok**?
	Are you crying when Daan is making fun of him?
nonword-nonword	Huil jij als Daan met hem **glot-glok**?
	Are you crying when Daan glot him?
real word-nonword	Weet jij of Kim vaker **breit-breik**?
	Do you know whether Kim knits regularly?
nonword-nonword	Weet jij of Kim vaker **gneit-gneik**?
	Do you know whether Kim gneit regularly?

lexical condition	nominal context
real word-nonword	Zie jij nog wel eens een **fuut-fuuk**?
	Do you now and then see a grebe?
nonword-nonword	Zie jij nog wel eens een **bluut-bluuk**?
	Do you now and then see a bluut?
real word-nonword	Wil jij een vloer van **hout-houk**?
	Do you want a floor made from wood?
nonword-nonword	Wil jij een vloer van **snout-snouk**?
	Do you want a floor made from snout?
real word-nonword	Huur jij een huis in die **straat-straak**?
	Do you rent a house in that street?
nonword-nonword	Huur jij een huis in die **klaat-klaak**?
	Do you rent a house in that klaat?
real word-nonword	Zie jij een map op die **boot-book**?
	Do you see a map on that boat?
nonword-nonword	Zie jij een map op die **zoot-zook**?
	Do you see a map on that zoot?

Materials for Experiments 6.1 & 6.2

The experimental stimuli are listed per item set and translations are given in italics.

Monosyllabic item set

lexical condition	verbal context (VP)
real word-nonword	Vraag jij of Jan morgen **vist-visk**? *Are you asking whether Jan fishes tomorrow*?
nonword-nonword	Vraag jij of Jan morgen **bist-bisk**? *Are you asking whether Jan bist tomorrow*?
real word-nonword	Vraag jij of Jan morgen **wast-wask**? *Are you asking whether Jan washes tomorrow*?
nonword-nonword	Vraag jij of Jan morgen **nast-nask**? *Are you asking whether Jan nast tomorrow*?
real word-nonword	Hoor jij wanneer Sander **praat-praak**? *Can you hear when Sander is talking*?
nonword-nonword	Hoor jij wanneer Sander **snaat-snaak**? *Can you hear when Sander snaat*?
real word-nonword	Hoor jij wanneer Sander **schiet-schiek**? *Can you hear when Sander is shooting*?
nonword-nonword	Hoor jij wanneer Sander **smiet-smiek**? *Can you hear when Sander smiet*?
real word-nonword	Lach jij als Bas haar **bijt-bijk**? *Are you laughing when Bas bites her*?
nonword-nonword	Lach jij als Bas haar **zweit-zweik**? *Are you laughing when Bas zweit her*?

Materials for Experiments 6.1 & 6.2 (continued)

Monosyllabic item set

lexical condition	nominal context
real word-nonword	Zie jij nog wel eens een **plaat-plaak**? *Do you now and then see a record*?
nonword-nonword	Zie jij nog wel eens een **snaat-snaak**? *Do you now and then see a snaat*?
real word-nonword	Hebben wij nog genoeg **gist-gisk**? *Do we still have enough yeast*?
nonword-nonword	Hebben wij nog genoeg **bist-bisk**? *Do we still have enough bist*?
real word-nonword	Hebben wij nog genoeg **riet-riek**? *Do we still have enough reeds*?
nonword-nonword	Hebben wij nog genoeg **smiet-smiek**? *Do we still have enough smiet*?
real word-nonword	Schilder jij nu een **geit-geik**? *Are you now drawing a goat*?
nonword-nonword	Schilder jij nu een **zweit-zweik**? *Are you now drawing a zweit*?
real word-nonword	Hebben wij soep voor de **gast-gask**? *Do we have soup for the guest*?
nonword-nonword	Hebben wij soep voor de **nast-nask**? *Do we have soup for the nast*?

Materials for Experiments 6.1 & 6.2 (continued)

Bisyllabic item set

lexical condition	verbal context (VP)
real word-nonword	Lach jij als Bas haar **verslaat-verslaak**? *Are you laughing when John beats her up*?
nonword-nonword	Lach jij als Bas haar **mejaat-mejaak**? *Are you laughing when John mejaak her*?
real word-nonword	Snap jij waarom Loes de school **bezet-bezek**? *Do you get why Loes occupies the school*?
nonword-nonword	Snap jij waarom Loes de school **lidwet-lidwek**? *Do you get why Loes lidwet the school*?
real word-nonword	Snap jij waarom Loes de school **verlaat-verlaak**? *Do you get why Loes leaves the school*?
nonword-nonword	Snap jij waarom Loes de school **beglaat-beglaak**? *Do you get why Loes beglaat the school*?
real word-nonword	Zeg eens waarom jij de wijn **verhit-verhik**? *Tell me why you are heating up the wine*?
nonword-nonword	Zeg eens waarom jij de wijn **lamit-lamik**? *Tell me why you lamit the wine*?
real word-nonword	Zeg eens waarom jij de wijn **verbiedt-verbiek**? *Tell me why you do not allow the wine*?
nonword-nonword	Zeg eens waarom jij de wijn **mediet-mediek**? *Tell me why you mediet the wine*?

Materials for Experiments 6.1 & 6.2 (continued)

Bisyllabic item set

lexical condition	nominal context (NP)
real word	Zie jij nog wel eens een **piraat-piraak**?
	Do you now and then see a pirate?
nonword	Zie jij nog wel eens een **mejaat-mejaak**?
	Do you now and then see a mejaat?
real word-nonword	Hebben wij soep voor de **soldaat-soldaak**?
	Do we have soup for the soldier?
nonword-nonword	Hebben wij soep voor de **beglaat-beglaak**?
	Do we have soup for the beglaat?
real word-nonword	Schilder jij nu een **buffet-buffek**?
	Are you now drawing a buffet?
nonword-nonword	Schilder jij nu een **lidwet-lidwek**?
	Are you now drawing a lidwet?
real word-nonword	Hoe beschrijf jij een **bandiet-bandiek**?
	How do you describe a bandit?
nonword-nonword	Hoe beschrijf jij een **mediet-mediek**?
	How do you describe a mediet?
real word-nonword	Hoe beschrijf jij een **gebit-gebik**?
	How do you describe dentures?
nonword-nonword	Hoe beschrijf jij een **lamit-lamik**?
	How do you describe a lamit?

Samenvatting

In dit proefschrift stonden twee thema's op het onderzoeksgebied van de gesproken taalverwerking centraal: segmentatie en foneem-beslissingen. Beide thema's werden vanuit een morfologisch perspectief benaderd. In twee reeksen van experimenten probeerde ik de relatie tussen morfolgie en vroege taalverwerkingsprocessen te onderzoeken. Aangezien morfologie de verbinding beschrijft tussen vorm en betekenis zou morfologie belangrijke informatie kunnen toevoegen aan alle drie de niveaus van het herkenningssysteem. Een centrale vraag hierbij is in welke vorm morfologisch complexe woorden zijn opgeslagen: zijn deze complexe woorden opgeslagen als geheel of in losse morfemen?

Morfologie en het segmentatie probleem

Het doel van het onderzoek in het eerste deel van dit proefschrift was een beter inzicht te krijgen in de karakteristieken van een specifiek segmentatie-principe: *The Possible Word Constraint* (PWC; Norris et al., 1997). Deze constraint zegt dat luisteraars het binnenkomende spraaksignaal zodanig segmenteren dat er geen onmogelijke woorden overblijven. Als het herkenningssysteem de twee Duitse woorden *Fall streiten* in de zin *Man kann über den Fall streiten* tegenkomt worden naast de woorden *Fall* en *streiten* ook de woorden *falsch* (fout) en *reiten* (paardrijden) tijdelijk geactiveerd. Echter, als het systeem de woorden *falsch* en *reiten* selecteert betekent dat dat de klank [t] overblijft. Een alleenstaande medeklinker zoals de [t] kan in het Duits geen mogelijk woord vormen (net zoals in alle andere Europese talen). De PWC zorgt ervoor dat het systeem de activatie van de woorden *falsch* en *reiten* onderdrukt zodat er geen alleenstaande medeklinker over blijft. De PWC maakt gebruik van verschillende cues, zoals de metrische structuur, fonotactische constraints en akoestische informatie, om de locatie van een mogelijke woordgrens te bepalen. Hoewel deze cues variëren tussen talen blijkt de PWC universeel te zijn (McQueen et al., 2001; Norris et al., 2001; Cutler et al., submitted). De PWC in zijn huidige versie zorgt ervoor dat de activatie van die woorden wordt verminderd, die een reeks fonemen zonder een klinker overlaten (bv. *wach* 'wakker' in *schwach* 'zwak').

De PWC en morfologie

In Europese talen waarin één enkele medeklinker een morfeem kan zijn (bv. de *-t* in de Nederlandse werkwoordsvorm *loopt*) heeft de PWC interessante consequenties voor de verwerking van vormen met inflectie. Als men de geïnflecteerde vorm *loopt* in de frase *zij loopt* hoort, zorgt de PWC ervoor dat de activatie van de om herkenning strijdende vorm *loop* verminderd wordt. De [t] tussen de *p* van *loop* en het eind van de frase vormt namelijk geen mogelijk woord. De luisteraar herkent dus zonder moeite het complexe woord *loopt*. Dit houdt in dat morfologisch complexe woorden zoals *loopt* als geheel moeten zijn opgeslagen in het lexicon. Als geïnflecteerde vormen in losse morfemen zijn opgeslagen (bv. *loop+t*) betekent dit, dat zowel de stam *loop-* als het inflexiemorfeem *-t* moeten worden geactiveerd om de vorm *loopt* succesvol te herkennen. Om te voorkomen dat de PWC de activatie van de vorm *loop* vermindert, met als gevolg dat de vorm *loopt* (opgebouwd uit *loop-* en *-t*) niet kan worden herkend, moet de PWC een uitzondering maken voor alleenstaande medeklinkers die een morfeem vormen (zoals de *-t*). In dat geval moet de PWC dus gevoelig zijn voor morfologische informatie.

Word-spotting experimenten

In de eerste serie experimenten werd de *word-spotting*-taak gebruikt om deze kwestie te onderzoeken. In experiment 2.1 waren monosyllabische woorden ingebed in drie verschillende soorten onzinwoorden. De eerste soort onzinwoorden was zodanig geconstrueerd dat de detectie van het ingebedde woord relatief makkelijk zou moeten zijn omdat de context die volgde op het ingebedde woord een lettergreep en dus een mogelijk woord in het Nederlands was (bv. *deur* in *deurtach*; de *lettergreep*-conditie). De tweede conditie echter was bedoeld om het spotten van de ingebedde woorden moelijker te maken door de context die volgde te laten bestaan uit één enkele medeklinker (bv. *deur* in *deurp*; de *medeklinker*-conditie). In de derde conditie was de medeklinker die op het ingebedde woord volgde een Nederlands inflectiemorfeem (de *morfeem*-conditie). De medeklinker was of het werkwoordsmorfeem *-t* (bv. *deur* in *deurt* [de stop-item-set]) of het meervoudsmorfeem *-s* (bv. *duim* in *duims* [de fricatief-item-set]). Het idee was dat het spotten van *deur* in *deurt* net zo makkelijk zou moeten zijn als het spotten van *deur* in *deurtach* als de PWC gevoelig is voor morfologische informatie. Als de PWC niet gevoelig is voor morfologische informatie zou het spotten van *deur* in *deurt* net zo moelijk moeten zijn als in *deurp*. Dit experiment repliceerde eerdere bevindingen in bijvoorbeeld het Engels, waar de reactie tijden (RT) sneller waren wanneer de volgende context een lettergreep was dan wanneer de volgende context één enkele medeklinker was (Norris et al., 1997). Echter, er waren geen significante verschillen tussen de *morfeem*-conditie en de twee andere condities. Dit werd grotendeels veroorzaakt door twee artefacten. Ten eerste werd de kracht van het experiment waarschijnlijk verminderd door het kleine aantal items (30 in totaal). Ten tweede waren de waarschijnlijkheden van de klankcombinaties aan het eind van de onzin woorden in de *medeklinker*-conditie van één item set (bv. *duim* in *duimf*) zo laag dat de taak op basis van alleen deze informatie kon worden voltooid. Aangezien de klanken *m* en *f* bijna nooit achter elkaar voorkomen in het Nederlands (Lugt, 1999, 2001), is het in dat geval makkelijk voor de luisteraar om de korrekte woordgrens te vinden.

Daarom werden de twee verschillende item-sets apart onderzocht in Experiment 3.1 zodat meer experimentele stimuli konden worden gebruikt. De monosyllabische items uit Experiment 2.1 werden aangevuld met bisyllabische zelfstandige naamwoorden. Bovendien waren de waarschijnlijkheden van de klankcombinaties aan het eind van de onzin woorden in de *fricatieve item-set* nu hoger dan in het eerste experiment. Het doel van het tweede *word-spotting*-experiment was een duidelijk resultaat te krijgen van de *morfeem* conditie. Maar in plaats daarvan werd een omgekeerd PWC effect gevonden. Dit betekent dat woorden sneller werden gedetecteerd in de *medeklinker*-conditie dan in de *lettergreep*-conditie. Dus de conditie die volgens de voorspellingen het langzaamst had moeten zijn leverde de snelste RTs op. Hoewel de RTs in de *morfeem*-conditie niet verschilden van die in de *medeklinker*-conditie, was het onmogelijk om dit effect te interpreteren, aangezien de onverwachte resultaten in de *lettergreep*-conditie aangaven dat er een artefact in de stimuli aanwezig was. Een controle-experiment onderzocht of de metrische verschillen tussen de condities het omgekeerde effect veroorzaakten, maar hetzelfde omgekeerde patroon kwam tevoorschijn.

Positieve correlaties tussen de RTs en de lengtes van de volgende contexten gaven aan dat luisteraars niet zo snel mogelijk reageerden maar wachtten tot het eind van het hele woord alvorens een reactie te geven. De twee volgende experimenten (3.3 en 3.4) introduceerden ieder een nieuwe factor om een verklaring te vinden voor het feit dat luisteraars wachtten tot het eind van een onzinwoord. De hypothese was dat luisteraars de ingebedde woorden al hadden geïdentificeerd voordat ze de rest van de context, die hun prestatie volgens de voorspellingen had moeten beïnvloeden, hadden gehoord. Er waren twee mogelijke oorzaken voor deze vroege identificatie: de vroege uniekheidspunten van de bisyllabische woorden en de vaste positie van de ingebedde target woorden, namelijk aan het begin van de onzinwoorden. Dus luisteraars konden ervan uitgaan dat het begin van de ingebedde woorden altijd samenviel met het begin van de langere onzin woorden. Beide factoren zouden de *word-spotting* taak eenvoudiger maken met als gevolg dat luisteraars als het ware de luxe hadden om te wachten tot het einde van het gehele onzinwoord alvorens te reageren.

In Experiment 3.3 werden items zo geconstrueerd dat targetwoorden zich niet alleen maar aan het begin van onzinwoorden bevonden (bv. *deur* in *deurp*) maar ook aan het eind van onzinwoorden (bv. *lepel* in *blepel*). Het feit dat luisteraars nu niet meer konden aannemen dat de targets zich altijd in dezelfde positie bevonden, zou de identificatie van de cruciale woorden (ingebed aan het begin) moeten vertragen. Daardoor zou de volgende context in staat zijn de prestatie van de luisteraars te beïnvloeden. De resultaten lieten zien dat aan-het-eind-ingebedde targetwoorden een robust PWC-effect opleverden (snellere RTs voor *lepel* in *kulepel* dan in *blepel*). Er werd echter opnieuw een omgekeerd effect gevonden voor aan-het-begin-ingebedde targetwoorden (snellere RTs voor *forel* in *forelf* dan in *forelfuil*).

In Experiment 3.4 werd, in tegenstelling tot in vorige experimenten, de lengte van de volgende contexten in de drie condities gelijk gehouden. Als het PWC-effect simpelweg overschaduwd werd door de strategie van de luisteraars om te wachten tot het eind van de onzinwoorden, zou dit kunnen helpen het PWC-effect te laten verschijnen. En inderdaad, de resultaten lieten een effect in de door de PWC voorspelde richting zien: targets werden sneller gespot in de *lettergreep*-conditie dan in de andere twee condities. Echter, positieve correlaties tussen de RT en de context-lengte gaven opnieuw aan dat het niet de factor context-*type* (*lettergreep, medeklinker* of *morfologisch*) was, die de RTs bepaalde maar de factor context-*lengte*. Het effect in het laatste experiment was in de verwachte richting simpelweg omdat nu (in tegenstelling tot in andere experimenten) de gemiddelde context-lengte in de *lettergreep*-conditie (de snelste conditie) significant korter was dan in de andere twee condities. Hieruit blijkt dat de poging om de context-lengte constant te houden niet was gelukt.

Uit deze serie experimenten blijkt dat de oorspronkelijke vraag over de rol van morfologie voor de PWC niet in het Nederlands beantwoord kan worden.

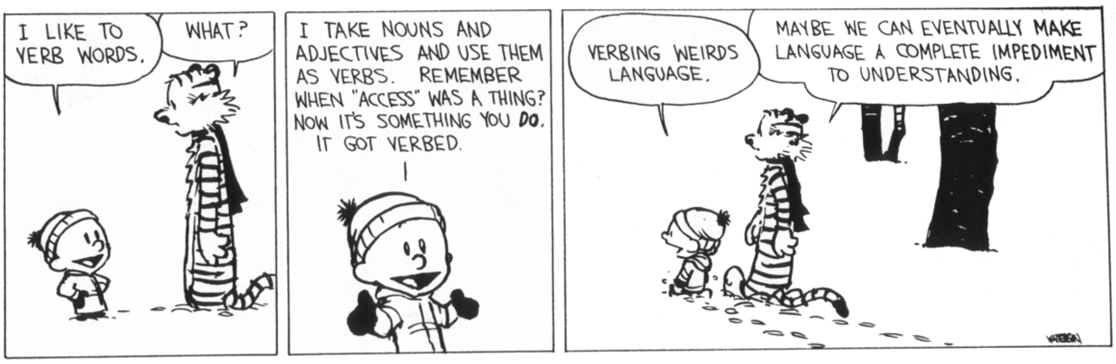
Foneem beslissingen en morfologische informatie

Foneem beslissingen over ambigue of vervormde klanken zijn gevoelig voor invloeden van verschillende vormen van informatie. Hiervan zijn de lexicale en de sententiële/**semantische** invloeden het meest beschreven in de literatuur. Er is echter minder aandacht besteed aan sententiële/**syntactische** invloeden op foneem beslissingen. Tot nu toe heeft geen enkele studie onderzocht of morfologische informatie en sententiële/syntactische informatie elkaar beïnvloeden tijdens het maken van een foneem beslissing. In het tweede deel van dit proefschrift probeerde ik de volgende vraag te beantwoorden: worden de responsies van luisteraars beïnvloed door morfologische informatie wanneer zij expliciet gevraagd worden te kiezen tot welke foneem catgorie een ambigue klank behoort. In zeven foneem categorisatie experimenten kregen luisteraars zinnen te horen die of een geïnflecteerd werkwoord bevatten (bv. **gaat** in *Vraag jij of Jan morgen gaat?*; de *werkwoord* context) of een zelfstandig naamwoord (bv. **plaat** in *Zie jij nog wel eens een plaat?*; de *naamwoord* context). Zowel de werkwoordsvormen als de zelfstandige naamwoorden eindigden met het foneem /t/, die in het eerste geval een morfeem is maar in het tweede niet. De syntactische structuur van de draagzin bepaalde de syntactische categorie van het laatste woord (een geïnflecteerd werkwoord of een zelfstandig naamwoord). Voor ieder van deze zinsfinale woorden werd een continuüm geconstrueerd waarbij het laatste foneem /t/ in stapjes werd veranderd in een /k/. Het ene uiteinde van de reeks vormde dus een woord (bv. *gaat*) en het andere uiteinde een onzinwoord (bv. *gaak*). De stapjes in het midden van de reeks waren ambigu tussen /t/ en /k/. Luisteraars werden gevraagd om de laatste klank van het laatste woord in de zin te categoriseren als een /t/ of een /k/. In dergelijke experimenten verwacht je allereerst een lexicaal effect (Ganong, 1980): luisteraars hebben de voorkeur de ambigue klank te identificeren als het foneem dat resulteert in een woord. De cruciale vraag is of er een verschil is tussen de categorisatie van de ambigue klank in de *gaat-gaak* reeks en de *plaat-plaak* reeks. In het eerste geval (de *werkwoord*-context) voorspelt de voorafgaande zinscontext (*Vraag jij of Jan morgen ...*) een werkwoord in de derde persoon enkelvoud en daarom het morfeem /t/ aan het eind van de zin. In het tweede geval (de *naamwoord*-context) voorspelt de voorafgaande zinscontext (*Zie jij nog wel eens een ...*) een zelfstandig naamwoord, maar echter niet het foneem waarop dit naamwoord eindigt. Om mogelijke verschillende resultaten voor geïnflecteerde werkwoorden en zelfstandige naamwoorden toe te kennen aan de verwerking van de voorafgaande zinscontext werden dezelfde woorden ook in isolatie aangeboden. Daarnaast werden de bestaande woorden vervangen door pseudowoorden waarin een bestaand woord ingebed was (**vla** in **vlaat**: *Vraag jij of Jan morgen vlaat?* en *Zie jij nog wel eens een vlaat?*). De vraag is of de prestatie van de luisteraars beïnvloed wordt door het feit dat de decompositie van het pseudowoord *vlaat* (in de onderdelen *vla* en *-t*) alleen zinvol is in de zinscontext die een zinsfinaal werkwoord voorspelt.

In de resultaten werd er naast het verwachte lexicale effect in beide contexten een additief effect gevonden: luisteraars waren geneigd de ambigue klanken vaker als /t/ te categoriseren in de *werkwoord* context dan in de *naamwoord* context. Deze neiging was het sterkst in de snelle responsies van de luisteraars.

Er werd echter, tegen de verwachtingen in, geen lexicaal effect gevonden voor de woorden in isolatie. Dit maakte het moelijk om het effect voor woorden in zinscontext te interpreteren. In tegenstelling tot de bestaande woorden was er bij de pseudowoorden geen lexicaal effect en geen zinscontext-effect. In geen van de twee contexten (*werkwoord*- en *naamwoord*-context) waren luisteraars geneigd meer /t/ responsies te geven. Dit geeft aan dat pseudowoorden zelfs in een context die een morfemische *-t* voorspelt (*werkwoord*-context) niet worden gedecomposeerd in morfemen. Uit deze experimenten bleek dus dat de luisteraars meer /t/ responsies gaven wanneer deze responsies resulteerden in een bestaand woord en wanneer de context een morfemische *-t* voorspelde. In deze experimenten was de voorspelbaarheid van de zinsfinale woorden echter heel groot doordat steeds dezelfde zinnen opnieuw werden aangeboden. Experiment 6.1 werd geconstrueert om te kijken of een lagere voorspelbaarheid van het materiaal dezelfde resultaten zou opleveren. In dit experiment kregen luisteraars meer verschillende zinscontexten en meer verschillende zinsfinale woorden aangeboden dan in de hiervoor beschreven experimenten. Pseudowoorden werden in dit experiment weggelaten aangezien ze in het voorafgaande experiment geen effecten opleverden. De nieuwe resultaten lieten, in tegenstelling tot eerdere resultaten, zien dat er geen lexicaal effect was voor werkwoorden in zinscontext. Er was echter wel een lexicaal effect voor zelfstandig naamwoorden in zinscontext. Wanneer deze woorden in isolatie werden aangeboden, vertoonden zowel werkwoorden als zelfstandige naamwoorden een lexicaal effect. Dit lexicaal effect was zwakker voor de geïnflecteerde werkwoorden dan voor de zelfstandige naamwoorden. Dit kwam waarschijnlijk door het feit dat normaal gesproken zelfstandige naamwoorden wel in isolatie kunnen voorkomen maar geïnflecteerde werkwoorden niet. Het lexicaal effect van de zelfstandige naamwoorden in isolatie en in zinscontext was van vergelijkbare grootte. Deze laatste overeenkomst suggereerde dat de categorisatie van ambigue klanken aan het eind van zelfstandige naamwoorden alleen door lexicale informatie werd beïnvloed. De verschillende resultaten voor werkwoorden in zinscontext en in isolatie werden toegekend aan een dissociatie van zinsverwerking en fonetische verwerking. Aangezien het laatste foneem zich in de *werkwoord*-context in een morfemische positie bevond waren luisteraars waarschijnlijk in staat om deze laatste klank onafhankelijk van de rest van de zin te categoriseren. Deze dissociate van twee verwerkingsprocessen was alleen mogelijk als het laatste foneem zich in morfeem positie bevond.

De resultaten van beide reeksen experimenten konden geen duidelijk antwoord geven op de vraag wat voor een rol morfologie speelt in de verwerking van gesproken taal. Hopelijk zal deze kwestie toekomstige onderzoekers blijven inspireren om zo uiteindelijk tot een volledigere beschrijving van het taalverwerkingssysteem te komen.



Curriculum Vitae

I was born in Köln, Germany, on July 12, 1971. After my graduation from the Albert-Schweitzer Gymnasium in 1990 in Hürth, I started studying German literature and Italian at the University of Cologne, Germany. After one year I decided to switch the focus of my studies. I moved to Düsseldorf where I studied Linguistics, English and Neurology at the Heinrich-Heine-Universität Düsseldorf, Germany. From October 1994 to March 1995 I spent half a year in Essex where I took courses in Linguistics at the University of Colchester, Great Britain. In 1997 I received my MA at the University of Düsseldorf, Germany. After that I worked as a research associate at the research project "Representation and processing of inflectional elements" at the linguistics department of the Heinrich-Heine-Universität Düsseldorf, Germany. In 1998 I was awarded a scholarship from the German *Max Planck Gesellschaft* to write my Ph.D. thesis at the *Max Planck Institute for Psycholinguistics* in Nijmegen, The Netherlands. From October 2000 to March 2001 I was a research associate at the *Max Planck Institute of Cognitive Neuroscience* in Leipzig, Germany.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing
   Miranda van Turennout
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography
   Niels Schiller
3. Lexical access in the production of ellipsis and pronouns
   Bernadette Schmitt
4. The open-/closed-class distinction in spoken-word recognition
   Alette Haveman
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach
   Kay Behnke
6. Gesture and speech production
   Jan-Peter de Ruiter
7. Comparative intonational phonology: English and German
   Esther Grabe
8. Finiteness in adult and child German
   Ingeborg Lasser
9. Language input for word discovery
   Joost van de Weijer
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe
    James Essegbey
11. Producing past and plural inflections
    Dirk Janssen
12. Valence and transitivity in Saliba, an Austronesian language of Papua New Guinea
    Anna Margetts
13. From speech to words
    Arie van der Lugt

14. Simple and complex verbs in Jaminjung: A study of event categorization in an Australian language
Eva Schulze-Berndt
15. Interpreting indefinities: An experimental study of children's language comprehension
Irene Krämer
16. Language-specific listening: The case of phonetic sequences
Andrea Weber
17. Moving eyes and naming objects
Femke van der Meulen
18. Analogy in morphology
Andrea Krott
19. Morphology in speech comprehension
Kerstin Mauth