Perceptual relevance of prevoicing in Dutch

ISBN: 90-76203-14-8

Cover design: Linda van den Akker, Inge Doehring
Cover illustration: Petra van Alphen
Printed and bound by Ponsen & Looijen bv, Wageningen

Perceptual relevance of prevoicing in Dutch

een wetenschappelijke proeve
op het gebied van de Sociale Wetenschappen

Proefschrift
ter verkrijging van de graad van doctor
aan de Katholieke Universiteit Nijmegen,
op gezag van de Rector Magnificus Prof. dr. C.W.P.M. Blom,
volgens besluit van het College van Decanen
in het openbaar te verdedigen
op maandag 2 februari 2004
des namiddags om 1:30 uur precies

door
Petra Martine van Alphen
geboren op 9 februari 1975 te Almelo

Promotor: Prof. dr. A. Cutler

Copromotores: Dr. J.M. McQueen (MPI)
Dr. R. Smits (MPI)

Manuscriptcommissie: Prof. dr. R. Schreuder
Prof. dr. S.G. Nooteboom (Universiteit Utrecht)
Dr. J. Andruski (Wayne State University, Detroit, USA)

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

Dag ventje
met de fiets
op de vaas
met de bloem
ploem ploem

(Fragment uit *Marc groet 's morgens de dingen,* Paul van Ostaijen)

VOORWOORD

Meer dan vier jaar lang heb ik me bezig gehouden met de spraakklanken /b/ en /d/, en met hun stemloze zusjes /p/ en /t/. Maar mijn PhD-tijd bestond uit veel meer dan *bupudutu*. Zoals het spraaksignaal bestaat uit verschillende akoestische *cues* die al variërend over tijd en frequentie gezamenlijk de betekenis van een uiting bepalen, zo zijn het vooral de verschillende mensen die op verschillende momenten en plekken deel uit maakten van mijn dagelijkse leven die mijn PhD-tijd betekenis gaven.

Ik begon mijn PhD-project in de *Comprehension Group* in januari 1999, nadat ik al enigszins was opgewarmd als stagiaire en onderzoeksassistent. Ik deelde een kamer met Arie en Nicole, die al snel mijn *kemmetjes* werden en mijn PhD-tijd een geweldige start gaven. Terwijl Arie peentjes zat te zweten op de laatste versie van zijn boekje verhuisden Nicole en ik naar de kamer aan de overkant, waar Kerstin ons al snel vergezelde en zo de *KPN* compleet was. We hadden een onvergetelijke tijd vol MaxKrant-activiteiten, wilde spinazie-avonturen, vliegende pingu's en serieuze gesprekken. Samen met Kerstin ontdekte ik het flamencodansen en terwijl Nicole haar kata's oefende, klapten wij onze buleria. Vanaf januari 2001 bracht ik 5 maanden door in het *Speech Lab* van Joanne Miller, waar ik hartelijk werd ontvangen. De *Macroom* werd speciaal omgetoverd tot *Petra's Office*, maar door de afwezigheid van ramen vergat ik soms hoe blauw de lucht buiten was. Ik was erg blij met Michèle die me zowel in het lab als daarbuiten wegwijs maakte en met wie ik vele dagen heb genoten van de *Boston blue sky*. In juni kwam ik weer terug, net op tijd om bij de verdediging van Andrea te zijn, met wie ik nog een paar leuke conferentiereisjes zou gaan maken. In die zomer organiseerde ik samen met de andere *Tubbs-meiden* de derde TuBBs-zomerschool, hetgeen een groot succes werd. Daarna begon ik dapper aan mijn akoestische metingen en ging vaak even mijn benen strekken. Zo belandde ik bij Simone en Joana, waar ik altijd plaats kon nemen op het zadel en even mijn verhaal kwijt kon, of bij Dannie, waar ik altijd veel te lang bleef kletsen en er altijd wel wat te lachen viel, of achter de piano, die altijd trouw stond te wachten in de kelder. In oktober verhuisde ik naar de oude kamer van Colin, zonder wiens aanwezigheid de kamer wel erg leeg voelde. Gelukkig werd ik al snel vergezeld door Keren, die vele kopjes thee inschonk en me bemoedigend toesprak terwijl ik mijn laatste experimenten aan het analyseren was en mijn bevindingen op papier probeerde te zetten. Het uitzicht, de thee en de kussentjes in de vensterbank werkten uitnodigend, want regelmatig kregen we bezoek: van Joana, mijn paranimf-wederhelft met wie ik dus veel te smoezen had, of van Martijn, die al mijn dropjes opat (en weer aanvulde), of van Simone, die ook toen ze al in Pittsburgh zat regelmatig even langs kwam, maar vooral van Dannie en Kerstin. Zij waren samen met Keren altijd bereid om naar mijn ingewikkelde verhalen over *echte prinsen* en *ambigue bloemen* te luisteren, en hun input heeft me erg geholpen. Ik ben erg blij dat ik met Dannie alle laatste loodjes heb kunnen delen en dat Kerstin en Keren mijn paranimfen willen zijn. Ik wil jullie bedanken voor jullie steun en vriendschap. Zonder jullie was het schrijven van dit proefschrift een stuk minder leuk geweest.

Mijn promotor Anne wil ik bedanken voor het creëren van de stimulerende wetenschappelijke omgeving. Het is een luxe om zoveel geïnteresseerde en deskundige collega's te hebben en ik wil dan ook alle leden (inclusief oud-leden en gasten) van de *Comprehension Group* bedanken voor alle hulp, ideeën en gezelligheid. Verder wil ik de volgende mensen bedanken: Arie en Anne Pier voor het inspreken van alle materialen, Martijn voor de hulp met de d-primes en het lezen van mijn samenvatting, Harald voor het kweken en snoeien van de bomen, Joan voor het suggereren van het tweede discriminatie-experiment, Delphine voor het kritisch meedenken, Mirjam en Taehong voor hun commentaar op delen van dit proefschrift, Inge voor alle grafische hulp (het oog wil ook wat) en mijn mede-PhD-studenten voor alle dingen die we samen hebben ondernomen (de zomerscholen, de Plak-dansfeestjes, het zingen en dansen op de promotiefeestjes, de leesgroepjes, de *Ph-beers* en de lunches). Ook de mensen van de administratie, receptie en technische groep wil ik bedanken voor alle hulp en gezelligheid, en Jan, Hans en Pim voor hun vrolijke goedemorgen-praatjes.

In deze lijst van Max Planck-mensen die ik bedank ontbreken nog twee belangrijke namen, omdat ik nu eenmaal graag het beste tot het laatst bewaar. Mijn grootste dank gaat uit naar mijn co-promotors, James en Roel. Ik heb enorm geboft hen als begeleiders te hebben. Niet alleen heb ik ontzettend veel van ze geleerd, ze gaven me ook altijd het gevoel dat ik er niet alleen voor stond. James neemt altijd alle tijd voor me en heeft maar een half (Nederlands) woord nodig om te begrijpen wat ik bedoel. Iedere keer wist hij me weer gerust te stellen als ik door de *beren* en *peren* het patroon niet meer zag. Roel is erg goed in het stellen van de juiste vragen en zorgde er zo voor dat ik het grote geheel niet uit het oog verloor. Bovendien werkt zijn Hollandse nuchterheid vaak erg relativerend. Na meer dan vier jaar onderzoek kan ik, zonder enige statistische toetsing, concluderen dat James en Roel de *primary cues* zijn voor de wetenschappelijke betekenis van mijn PhD-tijd. De volgende stap is nu om te laten zien dat het ook zonder deze *primary cues* gaat lukken.

Tot slot wil ik mijn vrienden bedanken voor alle leuke avonden en weekenden, die er voor zorgden dat ik iedere keer met nieuwe energie aan de slag ging, en met name mijn hartsvriendin Vanessa met wie ik al bijna 20 jaar lief en leed kan delen. De laatste regel van dank is voor mijn ouders Maarten en Ottoline voor hun steun en het vertrouwen dat ze in me hebben. Ze hebben me altijd gestimuleerd om datgene te doen waar ik plezier in heb, en dat is tot nu toe aardig gelukt.

TABLE OF CONTENTS

INTRODUCTION

CHAPTER 1

La parole est moitié à celuy qui parle, moitié à celuy qui l'escoute; cettuy cy se doibt preparer à la recevoir, selon le bransle qu'elle prend.

Speech belongs half to the speaker, half to the listener;
the latter must prepare to receive it according to the motion it takes.

MICHEL DE MONTAIGNE, LES ESSAIS III, PARIS (1588)

Physically speaking, the speech signal is no more than compressions and expansions of the air. When a person speaks, she temporarily changes the pressure of the surrounding air. In these changes in air pressure the message of the speaker is encoded. In order to understand the meaning of this message, listeners have to know how to decode the air pressure changes. Only then can the pattern of speech sounds make sense. Since the speech signal is transitory in nature, this decoding should take place quickly. Listeners are indeed surprisingly good at rapidly analyzing the speech signal. Within a very short time listeners are able to extract the relevant information from the speech signal and match these acoustic properties onto higher level linguistic forms. It is very unlikely that these linguistic forms represent full sentences, since the number of possible sentences that a listener might hear is infinite, which makes it impossible to store them all. It is therefore generally assumed that the stored linguistic forms are representations of word-like units. After all, there are a finite number of different words which have to be recognized in order to recognize any possible sentence. As soon as the words of an utterance are recognized, different types of information become available, including semantic, syntactic and pragmatic information. With this information the listener can determine the syntactic and semantic relationships among the different words encoded in the speech stream and hence understand the meaning of the whole utterance.

The matching between the incoming speech signal and the stored lexical representations is however not as straightforward as one might think on the basis of the apparent ease with which listeners understand speech. One property of the speech signal is that it is highly variable. Many different kinds of factors can influence the acoustic realization of a particular speech sound. For example, the phoneme /s/ is not always produced in the same way. Different speakers will produce this sound differently, since all kinds of factors will influence the acoustics, including the sex and age of the speaker, the size and shape of the vocal tract, the dialect of the speaker, the speaking rate and the speech style. In addition to these speaker-dependent factors, the acoustic environment in which the speaker is situated also influences the acoustics. Furthermore, neighboring sounds have a relatively large impact on the phonetic realization of a speech sound. For example, the phoneme /s/ in a word such as *soep* (soup) sounds clearly different from the /s/ in *sip* (glum). This is because, when speakers produce a particular sound, such as /s/, they already prepare for the production of the following vowel, so the articulation of the /s/ is influenced by the following /u:/ or /ɪ/. This is called coarticulation. The effect of coarticulation also takes place across word boundaries, such that the realization of the /s/ in *soep* when it appears in the sentence *ik wil soep* (I want soup) is different from when it appears in the sentence *ik neem soep* (I take soup). And even if a word is produced several times by the same speaker in the same environment and in the same context, it will never be realized in exactly the same way. The fact that the speech signal is so highly variable makes the recognition of speech a complex process.

As a result of the variability of speech and the fact that there is only a limited phonetic space in which acoustic properties vary, some speech sounds show considerable phonetic overlap. Take for example the two phonemes /b/ and /p/. Although they are phonemically different, phonetically they are very similar. One of the acoustic cues which signal the difference between these two plosives in medial position, for example in the Dutch words *aanpakken* (to take) and *aanbakken* (to get burnt), is the closure duration. In order to produce a plosive, first all outgoing pathways are blocked, such that no air can get out. Immediately following the closure there is a buildup of pressure behind the constriction. The release of the constriction produces a noise burst. The closure duration is in general longer for voiceless plosives (e.g., /p/) than for voiced plosives (e.g., /b/). But the closure duration also changes with speaking rate. When a speaker speaks faster the closure will become shorter. Therefore it could be the case that the difference in the closure duration of two versions of the word *aanpakken* at two different speaking rates is in fact larger than the difference between the closure duration between the word *aanpakken* and the word *aanbakken* at a single speaking rate. Nevertheless, listeners must know that in the former case the large difference in closure duration does not signal a phonemic difference between /p/ and /b/, but that in both cases the word *aanpakken* is meant, while in the former case the small difference in closure duration signals a difference between /b/ and /p/ and that the speaker intended to say two different words. In other words, the listener should know which acoustic details are relevant for distinctions between phonemes and thus between words, and which acoustic details are not relevant for lexical distinctions. This thesis is mainly concerned with the way in which the speech recognition system deals with relevant and irrelevant acoustic detail as the variable speech signal is mapped onto the mental lexicon.

The relevance of different types of acoustic detail is closely connected to the way in which acoustic properties vary in natural speech. Acoustic cues which vary in a completely random way are obviously not very helpful for listeners. Thus, listeners should rely most strongly on those cues that systematically signal a distinction. They should therefore be sensitive to the distributional information of acoustic properties. Research in different areas of spoken language comprehension has shown that listeners are indeed sensitive to various types of distributional information. For example, listeners are sensitive to the different frequencies with which particular syntactic structures occur (e.g., Trueswell, Tanenhaus & Kello, 1993; Garnsey, Pearlmutter, Myers & Lotocky, 1997). Listeners also have knowledge about the frequency with which words occur. High frequency words are usually recognized faster than low frequency words (e.g., Soloman & Postman, 1952). For example, the Dutch high frequency word *boom* (tree) tends to be recognized faster than the Dutch low frequency word *biels* (sleeper). In addition to knowledge about the probability of the occurrence of a particular word in a particular syntactic structure, the probability that particular speech sounds occur after one another is also used by listeners in language processing (e.g., Vitevitch & Luce, 1999; van der Lugt, 2001). It thus seems that listeners have different kinds of distributional knowledge which can help them in the process of language comprehension. In order to make predictions about the way in which variation in a particular acoustic cue will influence the recognition of words it is therefore essential to first establish how this acoustic property varies in natural speech. This is exactly the purpose of Chapter 2 of this thesis.

Besides the fact that the speech signal is highly variable, the speech signal is also very rich in information. Each phonological distinction is signalled by the speaker by several different acoustic properties. In the right circumstances almost all these properties can function as perceptual cues to the identity of the speech sound. Not all cues are however of the same importance; some cues have a stronger influence on the phonemic percept than others and some cues can only carry weight when other cues are absent (Repp, 1982; Repp & Liberman, 1987). Furthermore, particular cues contribute to the identification of more than one linguistic unit (e.g., Smits, 2001). There is no one-to-one relationship between the acoustic properties and the stored mental representations. This makes the matching between speech input and stored representations a complex process. On the other hand, the existence of multiple acoustic properties for one particular linguistic unit is also one of the strengths of speech, since it makes speech recognition robust. If one acoustic property is missing (for example due to extraneous noise) there are usually other cues present that can be used instead such that perception does not suffer. Therefore, when the influence of one particular acoustic cue on perception is investigated it is important to examine the presence and strength of other acoustic cues. This is also one of the purposes of Chapter 2.

PREVOICING

The acoustic property under investigation in this thesis is prevoicing. Prevoicing is the presence of vocal fold vibration during the closure of a plosive. It plays an important role in the realization and perception of the phonological voicing distinction between Dutch voiced and voiceless plosives at the beginning of an utterance. In Dutch there are two voiced plosives, namely [b] and [d], and three voiceless plosives, namely [p], [t] and [k]. Note that the velar voiced plosive [g] rarely occurs in Dutch; it only appears in loan words, for example in the word *goal*. Therefore the research is this thesis will be concerned with the difference between the voiced plosives [b] and [d] and the voiceless plosives [p] and [t].

In most languages there is a phonological voicing distinction in plosives, but the way in which this distinction is implemented phonetically differs among languages. Voice Onset Time (VOT) plays an important role in this phonetic realization. VOT is the time between the onset of vocal fold vibration and the release of a plosive. In Dutch, voiced plosives in initial position are in general produced with a negative VOT (that is, the vocal folds start vibrating before the release of the closure), while voiceless plosives in initial position are in general produced with a slightly positive VOT which lies close to zero (that is, the vocal folds start vibrating soon after the release). I will use the term prevoicing to refer to the presence of voicing during the closure of a plosive. The main goal of this thesis was to examine the effects of prevoicing variation on lexical processing.

LEXICAL AND PRELEXICAL PROCESSING

Our mental lexicon contains representations of all the words we know. The number of stored lexical representations range from 10,000 to 100,000, depending on how one defines what a lexical representation is (for example, whether the plural form *eyes* is stored separately from it’s singular form or not). All these lexical representations can be activated as a result of the incoming speech signal. This activation process operates in cascade, meaning that while the acoustic realization of a word unfolds over time the activation patterns at the lexical level change continuously. This implies that at the beginning of an utterance multiple lexical candidates will be activated simultaneously, because the listener does not know what acoustic information will follow. Thus, upon hearing the initial part /su:/ of the word *soep*, not only the lexical representation of *soep* but also of *soepel*, *soesje* and *soeverein* will be activated. This assumption is central to the Cohort model (Marslen-Wilson & Welsh, 1978). As more acoustic information becomes available, particular lexical representations will be activated more strongly than others. When one of the lexical candidates has received significantly more activation than the other candidates, it will be recognized. Thus the degree of lexical activation reflects the goodness-of-fit between the lexical representations and the available acoustic evidence at that moment (e.g., McQueen, Dahan & Cutler, 2003). Research has shown that the degree of activation of lexical candidates is not only influenced by acoustic evidence but also by the degree of activation of the other lexical candidates (see McQueen, in press). Candidates which receive more activation than others suppress the activation of the other lexical candidates. In other words, lexical candidates compete with each other for activation. In this way, small differences in the acoustic signal can result in large differences in activation levels, such that one candidate has clearly been activated more strongly than the other candidates and can easily be recognized as the winner.

The matching process between the acoustic signal and lexical representations is a complex process due to the variability of the speech signal and the absence of a one-to-one relationship between acoustic properties and linguistic units. Some models of word recognition therefore assume that the acoustic signal is directly mapped onto lexical representations (e.g., Klatt, 1979; Goldinger, 1998). This means that all the acoustic properties of all words should be stored in the mental lexicon. In other words, the phonetic details of a particular speech sound which occurs in many words is then stored many times in different lexical representations. For example, the information that a voiced plosive such as /b/ has voicing during the closure should then be stored separately for each word which contains a /b/. This would result in a large amount of redundant information in the lexicon. To solve this problem of redundancy many models of word recognition including TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994; Norris, McQueen & Cutler, 2000) assume that there is an intermediate stage of processing. At this prelexical level the acoustic input is analyzed and relevant information is extracted from the speech signal and mapped onto abstract prelexical representations. What these representations are is still unclear. Different units have been proposed, varying from acoustic features (Stevens, 2002) to allophones (Luce, Goldinger, Auer & Vitevitch, 2000), phonemes (Foss & Blank, 1980; Nearey, 2001), semi-syllables (Massaro, 1987) and syllables (Mehler, 1981). In this thesis I refer to phonemic representations at the prelexical level when explaining the effects of prevoicing variation, but this does not imply that I want to claim that these are *the* prelexical units of perception.

A large body of research on the effects of fine acoustic detail on lexical activation has given more insight into the way in which the prelexical level operates, in particular what type of information is passed on to the lexical level and is therefore not normalized away at the prelexical level. In discussing the results of a number of these experiments I would like to make a distinction between experiments in which acoustic details have been manipulated in such a way that it has resulted in an unnatural use or combination of acoustic information, and experiments in which the investigated acoustic differences appear naturally in speech. Whereas the former group of experiments provide useful information about the sensitivity and tolerance of the recognition system towards mismatching acoustic information, the results of the latter group shed light on the way the system deals with acoustic variation which appears naturally.

Experiments in the first of these classes have examined the effects of mismatches in acoustic-phonetic information (e.g., Dahan, Magnuson, Tanenhaus & Hogan, 2001; Marslen-Wilson & Warren, 1994; McQueen, Norris & Cutler, 1999; Streeter & Nigro, 1979; Whalen, 1984, 1991). The materials in these studies were obtained by cross-splicing stretches of speech which originated from different words and nonwords. For example, Dahan et al. created three different versions of the word *net* by cross-splicing the first two phonemes of either the word *net*, or the word *neck* or the nonword *nep* onto the final phoneme of another token of the word *net*. In this way the formant transitions in the vowels of these three versions of the word *net* signalled different places of articulation of the following phoneme. Thus the coarticulatory information in the vowel was either in favor of the final consonant /t/, or it was in conflict with the final consonant /t/. The effects of these acoustic differences were tested in an eye-tracking experiment. During the experiment participants were seated in front of a computer screen with four different pictures, one of which displayed the target word (*net*). Participants heard spoken instructions to click on the *net* (in which the final word was one of the three cross-spliced versions). The results showed that participants were slower to fixate the target picture when the vowel of *net* was derived from the word *neck*. This suggests that the lexical representation of *neck* was also temporarily activated. This was confirmed by a second study in which one of the pictures on the screen displayed the competitor word *neck*. When the vowel of the word *net* was derived from the word *neck*, the proportion of fixations to the competitor word *neck* was temporarily higher than when the vowel was derived from the word *net*. This activation of the competitor word *neck* temporarily suppressed the activation of *net* and therefore the recognition of the target word *net* was delayed.

These effects show that lexical activation is sensitive to mismatching acoustic information. But this type of acoustic detail does not normally appear in natural speech. Except for the participants who took part in these experiments, no listener would ever encounter such mismatches in their daily conversations. There are, however, several experiments which have investigated the influence of natural acoustic variation on lexical activation (e.g., Spinelli, McQueen & Cutler, 2003; Gow & Gordon, 1995; Davis, Marslen-Wilson & Gaskell, 2002; Salverda, Dahan & McQueen, 2003). All these studies show that listeners are sensitive to small differences in the acoustic detail in the speech signal, for example to small durational differences. Salverda, Dahan & McQueen (2003) found that listeners are sensitive, for example, to differences in the duration of the syllable /hɑm/ in the Dutch words *ham* and *hamster.* When listeners heard a spliced version of the word *hamster* in which the first syllable was derived from a natural production of the word *ham* (and was thus longer), the proportion of fixations to the picture of a *ham* was higher than when the first syllable was derived from another token of *hamster* (and was thus shorter).

A study by Andruski, Blumstein and Burton (1993) focussed on another durational difference, namely the difference in the VOT in English initial plosives. They used a uni-modal associative priming task to explore the effects of VOT differences on lexical activation. The results indicated that the lexical representation of a word starting with a voiceless plosive, for example *king*, was activated less strongly upon hearing an auditory version of this word from which two-thirds of the original VOT had been removed than upon hearing a version with unaltered VOT. Similar findings were obtained by Utman, Blumstein and Burton (2000) using a uni-modal identity priming task. In English, degree of lexical activation is therefore affected by the duration of the positive VOT. In Chapters 3 and 4 of this thesis, several priming experiments are presented in which the effect on lexical activation of differences in negative VOT in Dutch was examined. Furthermore, Utman (1997) examined the effects of several acoustic properties on the recognition of English words ending with voiced plosives. One of these properties was the presence or absence of voicing during the closure of the final plosive. The results of a uni-modal associative priming experiment suggested that a lexical candidate with a final voiced plosive, for example *pig*, was activated less strongly upon hearing the altered version of *pig* without voicing during the closure than upon hearing the unaltered version of the word with voicing during the closure. In English, lexical activation is thus also affected by the deletion of voicing from the closure of final voiced plosives. In the present study, the effect of prevoicing deletion on the activation of Dutch words with initial voiced plosives was examined.

These effects of fine-grained acoustic information on lexical activation suggest that particular acoustic details are not normalized away at the prelexical level, but can influence the lexical level. Therefore, the representations at the prelexical level should somehow be capable of preserving these acoustic details. Most current models of word recognition therefore now assume that prelexical representations are activated in a graded fashion. That is, no concrete decisions (for example phoneme decisions) are made at the prelexical level, and this graded activation is passed on to the lexical level continuously. In this way, degree of lexical activation reflects the goodness-of fit with the speech signal.

In addition to models of spoken word recognition, there are also models which are primarily concerned with speech-sound perception. Some of these models, for example FUL (Lahiri & Reetz, 1999) assume that the prelexical level makes binary decisions. In FUL, a speech sound is either [+voice] or [-voice] and the match between these phonological features and the lexical representations can be threefold: either there is a match, or there is a mismatch or there is no mismatch. In the last case the absence of a particular feature does not help word recognition but does it does not hinder recognition either. On the basis of this account a speech sound is either voiced or voiceless. I will argue on the basis of the results of the experiments in this thesis that plosives are not simply voiced or voiceless, but that there are gradations of voicing. The probability that a plosive is voiced is determined by the combined evidence provided by all available acoustic cues, and as a result some tokens will be “less voiced” than others. This gradedness of activation at the prelexical level is then passed on to the lexical level. In turn, a particular lexical candidate starting with a voiced plosive can achieve different degrees of activation as a result of the amount of acoustic evidence in favor of the voiced plosive.

THE CURRENT STUDY

This thesis is organized as follows. In Chapter 2, the occurrence of prevoicing in initial Dutch voiced plosives was investigated in order to determine the range of natural variation in prevoicing. The influence of several factors on the frequency and duration of prevoicing was examined. Furthermore, the presence of several other potential acoustic cues was established and the relative strength of their influence on the perception of the voicing distinction were determined. Chapter 3, which was written as a separate book chapter, first summarizes the findings of Chapter 2 and then presents three experiments in which the effects of differences in the amount of prevoicing on word recognition were investigated. At the end of this chapter some of the results of Chapter 4 are briefly discussed. Chapter 4 describes three experiments which focussed on the influence of the existence of voiceless word competitors on the effects of prevoicing variation on lexical activation. Finally, in Chapter 5, the main findings of the thesis are summarized and tied together, leading to a discussion about the way in which the speech recognition system treats acoustic detail.

ACOUSTICAL AND PERCEPTUAL ANALYSIS OF THE VOICING DISTINCTION IN DUTCH INITIAL PLOSIVES: THE ROLE OF PREVOICING

CHAPTER 2

Petra M. van Alphen and Roel Smits (submitted), *Journal of Phonetics*

ABSTRACT

Three experiments investigated the voicing distinction in Dutch initial labial and alveolar plosives. The difference between voiced and voiceless Dutch plosives is generally described in terms of presence or absence of prevoicing (negative voice onset time). Experiment 2.1 showed, however, that prevoicing was absent in 25% of voiced plosive productions across 10 speakers. The production of prevoicing was influenced by place of articulation of the plosive, by whether the plosive occurred in a consonant cluster or not, and by speaker sex. Experiment 2.2 was a detailed acoustic analysis of the voicing distinction, which identified several acoustic correlates of voicing. Prevoicing appeared to be by far the best predictor. Perceptual classification data of Experiment 2.3 revealed that prevoicing was indeed the strongest cue that listeners use when classifying plosives as voiced or voiceless. In the cases where prevoicing was absent, other acoustic cues influenced classification, such that some of these tokens were still perceived as being voiced. These secondary cues were different for the two places of articulation. We discuss the paradox raised by these findings: although prevoicing is the most reliable cue to the voicing distinction for listeners, it is not reliably produced by speakers.

INTRODUCTION

In phonetic research, the term 'acoustic correlate' is often used to indicate an acoustic property which covaries with a phonemic distinction. A large body of phonetic research has been devoted to identifying such acoustic correlates for a number of phonetic distinctions. This research has shown that each phonemic distinction has several acoustic correlates. Subsequent perceptual experiments, employing synthetic stimuli in which one or more of these correlates were systematically varied, have shown that listeners are sensitive to many or all of these correlates when recognizing phonemes. An acoustic correlate which influences the perception of a phonemic distinction is often referred to as an 'acoustic cue' to that distinction.

The present study focuses on the voicing distinction in Dutch initial plosives, that is, the phonological distinction between [+ voice] and [-voice]. In particular, we aimed to identify the most important acoustic correlates of voicing in Dutch initial plosives, to establish which of these correlates are theoretically most reliable for recognizing voicing, and to determine which of the correlates are actually the strongest cues in listeners' categorizations. Our use of natural speech stimuli enabled us to study voicing perception when the full array of acoustic cues was present.

The phonological distinction between [+voice] and [-voice] in plosives has been one of the most intensively studied distinctions. Most languages contrast these two phonemic classes (which I will refer to as *voiced* and *voiceless* plosives), but the phonetic realization of this phonological distinction varies among languages. The moment that the vocal folds start vibrating relative to the moment of the release of the closure, the so called Voice Onset Time (VOT), plays an important role in these different acoustic realization. The notion of VOT was introduced by Lisker and Abramson (1964) who measured the VOT of plosive production in eleven languages. They concluded that, across languages, three different VOT categories could be distinguished. The first category of plosives had a negative VOT, that is, they were produced with voicing during the closure. The second category of plosives had a slightly positive VOT; these plosives were produced with little or no aspiration. The third class had a clear positive VOT; these plosives were produced with aspiration. Given these three VOT categories, any language could thus in principle employ a three-way voicing distinction There are, however only a few languages, for example Thai, which contrast these three voicing categories (fully voiced, voiceless unaspirated and voiceless aspirated) in plosives. Most languages have a two-way voicing distinction, which is implemented by two adjacent modes, one of which is associated with the phonologically voiced, and the other with the phonologically voiceless plosive. A study by Keating, Linker and Huffman (1983), in which 51 different languages were surveyed, showed that the voiceless unaspirated category is the most common category; it is used in almost all these languages. The two other categories, the fully voiced and voiceless aspirated category, appear equally often as the voicing category contrasting with the voiceless unaspirated category. Furthermore, Keating et al. (1983) observed that in many languages the use of these different VOT categories vary as a function of the position in a word at which the plosive occurs. In the present study we will focus on plosives in initial position of words spoken in isolation.

The way in which the voicing distinction is implemented phonetically is different in Dutch than in most other Germanic languages. While most Germanic languages such as Danish, English and German contrast voiceless unaspirated and voiceless aspirated plosives in initial position (Keating, 1984), Dutch does not. Dutch belongs to the group of languages, including for example Arabic, Bulgarian, French, Japanese, Polish, Russian and Spanish, which has a “traditional” voicing contrast (Keating, 1984; Lisker & Abramson, 1964). That is, the initial voiced plosives are produced with a negative VOT, which we will refer to as *prevoicing*, and the initial voiceless plosives are produced with little or no aspiration. In Dutch there are three voiceless plosive categories, namely [p], [t], and [k], but there are only two voiced plosive categories, namely [b] and [d]. The voiced velar plosive [g] is marginally present in Dutch as it only occurs in Dutch loan words (e.g., *goal*). Therefore, the voicing distinction in the velar plosives was not included in the present study.

Prevoicing is the production of vocal fold vibration during the closure phase of a plosive in initial position. Vocal fold vibration can only occur when certain physiological and aerodynamic conditions are met (van den Berg, 1958). First, the vocal folds must be properly adducted and tensed. Second, a sufficient transglottal pressure gradient is needed to result in enough positive airflow through the glottis to support vibration. The second condition is relatively hard to meet in the case of the closure of a plosive, since all outgoing pathways are closed. As a consequence, the air flowing through the glottis accumulates in the oral cavity, causing oral pressure to approach subglottal pressure (Ohala, 1983). This process will be delayed as the volume above the glottis increases. Therefore, expansion of the vocal tract volume will facilitate the production of voicing during closure. A part of the expansion can be achieved by active enlargement of the supraglottal cavity, namely by lowering the larynx, raising the soft palate, advancing the tongue root, or drawing the tongue dorsum and blade down (see Westbury, 1983). In addition to active enlargement, supraglottal volume can also be expanded passively due to the raised intraoral pressure, provided that the walls of the supraglottal cavity are lax (Rothenberg, 1968). Although it is difficult to differentiate changes in vocal tract size resulting from active and passive expansion (Westbury, 1983), it is generally assumed that both mechanisms play a role in the production of prevoicing.

The fact that extra articulatory movements are required to produce prevoicing makes the production of prevoicing relatively difficult. Several studies have shown that children who acquire languages that contrast voiced plosives with prevoicing and voiceless unaspirated plosives master the adult pattern later than children who acquire languages that contrast voiceless unaspirated and voiceless aspirated plosives (e.g., Allen, 1985; Kewley-Port & Preston, 1974; Konefal & Fokes, 1981; Macken & Barton, 1980). The late acquisition of prevoicing may be due to the relatively small vocal tract size in children (Rothman, Koenig & Lucero, 2002) and due to the complexity of the articulatory gestures which are demanded for the expansion of the vocal tract (Kewley-Port & Preston, 1974). It is important to bear in mind, however, that although the production of prevoicing is relatively difficult, it is not the case that it is simply too difficult, otherwise voiced plosives would not exist in such a considerable number of languages (Westbury & Keating, 1986).

Given the extra articulatory effort it takes to produce prevoicing, it is likely to be difficult for speakers to control the exact duration of prevoicing. Sometimes they may even fail to produce any prevoicing. Nevertheless, studies which have investigated the occurrence of prevoicing suggest that prevoicing is rarely absent in initial voiced plosives (Keating, Mikos & Ganong, 1981 on Polish; Yeni-Komshian, Caramazza & Preston, 1977 on Lebanese Arabic; Caramazza & Yeni-Komshian, 1974 on European French). Only one study, on Canadian French (Caramazza & Yeni-Komshian, 1974), has found a substantial degree of overlap between the VOT distributions of voiced and voiceless plosives; no less than 58% of the voiced tokens in that sample (N=90) were produced without prevoicing. Caramazza and Yeni-Komshian argued that in Canadian French the VOT values are shifting as a result of the influence of Canadian English. There are, however, no studies which systematically investigated the occurrence of prevoicing in Dutch. One of the goals of the present study was therefore to gain more insight into the way in which prevoicing varies in Dutch. Several factors were included which could have affected the occurrence of prevoicing and its duration.

The most complete study of the acoustics and perception of voicing distinction in Dutch plosives was conducted by Slis and Cohen (1969). Apart from prevoicing, they measured several additional acoustic properties that were known to play a role in the voicing contrast. They did not always describe, however, the details of their elicitation and measurement procedures. They also investigated the influence of several acoustic properties on the perception of voicing using synthetic speech. As in most studies in other languages, they varied each acoustic correlate separately (or maximally two at a time). In this way an overview was given of the way in which each of the acoustic properties influenced listeners' perception of voicing when all the other cues were kept constant. This analysis did not show, however, how all these acoustic properties vary together in natural speech nor which of the acoustic correlates are relied on most strongly by listeners. A full understanding of the phonetics of voicing requires an analysis of the variability in natural utterances of voiced and voiceless plosives, and an analysis of how listeners deal with that variability. The present study sought to provide such analyses. Because this is the first large-scaled study of the above issues in Dutch, we focused on voicing in plosives in initial position in words spoken in isolation. We note that sentence context may influence the phonetic realization of the voicing distinction (e.g., Lisker & Ambramson, 1964).

This study consists of three experiments. Experiment 2.1 was designed to investigate variation in the production of prevoicing in Dutch plosives and whether the presence or absence of prevoicing and the duration of prevoicing is influenced by a number of potentially relevant factors. Experiment 2.2 is a detailed acoustic analysis of the voicing distinction in Dutch plosives. Several acoustic properties in addition to prevoicing were measured and analyzed in order to find out which of these properties were correlates of the voicing distinction. Subsequently, classification tree analyses were used to indicate which of these acoustic correlates would serve as the most reliable cues for correct recognition of voicing. Experiment 3.3 investigated how the tokens of Experiment 2.2 were perceived by listeners and asked which of the acoustic properties identified in Experiment 2.2 are relied on most strongly by listeners when they identify plosives as voiced or voiceless. Together, the three experiments provide a detailed analysis of the production and perception of the voicing distinction in Dutch initial plosives with particular emphasis on the role of prevoicing.

EXPERIMENT 2.1

Although it is generally assumed that the presence of prevoicing is one of the major attributes of the voiced-voiceless distinction in Dutch plosives, there are to our knowledge no published studies that actually report acoustic measurements on prevoicing in Dutch other than the study conducted by Lisker and Abramson (1964) and the study by van Dommelen (1983). All Dutch tokens of /b/ and /d/ analyzed by Lisker and Abramson were produced with a voice lead. However, their measurements were based on the production of only one speaker. Furthermore, the way in which the speech was elicited was not described. The items may have been presented without fillers, in which case the speaker might have been aware of the type of distinction under investigation. This could have stimulated him to hyper-articulate, which may have resulted in more prevoicing than may normally occur in Dutch. The VOT values reported by van Dommelen showed that prevoicing was sometimes absent in initial voiced plosives, but these values were based on only a few different words.

The purpose of the first part of the present study was therefore to conduct a systematic and large-scale study of prevoicing variation in Dutch voiced plosives. In particular, we aimed to find out whether voiced initial plosives were consistently produced with prevoicing and whether the presence or absence of prevoicing and its duration varied as a function of several factors. The influence of the sex of the speaker was investigated as well as the influence of two segmental and two lexical factors: the place of articulation of the plosive (labial versus alveolar); the phoneme following the plosive (vowel versus consonant); the lexical status of the carrier stimulus (word versus nonword); and the competitor environment of the carrier stimulus, that is, whether changing the first voiced plosive into its voiceless counterpart resulted in a word or a nonword (competitor versus no competitor). The effect of two of these factors, namely the sex of the speaker and the place of articulation, were also investigated in an experiment on the occurrence of prevoicing in English (Smith, 1978). Note, however, that prevoicing is not important for the voicing distinction in English, since the difference between voiced and voiceless plosives is signalled by a difference in the positive VOT. To ensure robustness of our results, we recorded several speakers.

For two of the five factors, namely the sex of the speaker and the place of articulation of the plosive, we had clear predictions. The prediction for the speakers’ sex is based on differences in vocal tract size between men and women. The volume of the vocal tract is smaller in female than in male speakers (e.g., Stevens, 1998). Assuming equal volume velocity through the glottis, oral pressure will tend to rise more quickly in females than in males, which makes it harder to produce prevoicing. Smith (1978) indeed found that in English, prevoicing was less often produced by female speakers than by male speakers. In line with these findings, we therefore predicted a smaller proportion of prevoiced tokens in female speakers in comparison to male speakers.

As described earlier, one phenomenon that helps to maintain sufficient transglottal pressure is passive enlargement of the oral cavity due to raised intraoral pressure. For alveolar plosives, the pharyngeal walls and part of the soft palate can yield to expansion of the oral cavity, while for labial plosives these surfaces plus all of the tongue surface and parts of the cheek can participate in the expansion (Houde, 1968; Rothenberg, 1968). Furthermore, the freedom to actively expand the vocal tract through movements of the tongue body is expected to be smaller for /d/ than for /b/ since in the former case the tongue is already involved in maintaining the closure. The oral cavity can thus be expanded more during the production of labial plosives than during the production of alveolar plosives. Consequently, oral pressure tends to rise less quickly during the production of labial plosives than during the production of alveolar plosives. According to this account, the production of prevoicing was expected to be easier for labials than for alveolars. In line with this, the study by Smith (1978) showed that in English, place of articulation affected both the duration of prevoicing and the occurrence of prevoicing in the predicted direction. Therefore, Dutch labial plosives were also expected to be produced more often with prevoicing and with longer prevoicing than alveolar plosives.

The other segmental factor, namely the phoneme that followed the plosive, was included to test for possible differences between items in which the plosive was followed by a vowel and items in which the plosive was followed by a consonant. It is likely that the anticipatory coarticulation of the following phoneme affects vocal tract size and the degree to which the vocal tract size can be expanded. Smith (1978) for example found that the height of the following vowel had an influence on both the proportion of prevoiced tokens and on the duration of prevoicing. Although it is difficult to make detailed predictions for different consonants and vowels without the use of articulatory measurements, we included this factor in order to test whether the following phoneme influenced prevoicing production.

In addition to these two segmental factors, two lexical factors were included. It is often said that nonwords are hyper-articulated, although to our knowledge this claim has not been systematically investigated. Such hyper-articulation might cause plosives in nonwords to be produced more often with prevoicing and with longer prevoicing durations than plosives in words. Allen and Miller (2001) found that the lexical status of items had no effect on VOT duration in English voiceless plosives. The production of prevoicing is however more difficult than the production of aspiration. Therefore it might still be the case that prevoicing is more frequently omitted or shorter in words than in nonwords. For this reason, the lexical status of the item was included in the analysis. Furthermore, it is also possible that the lexical competitor environment of the carrier stimulus influences the production of prevoicing. Speakers might speak more carefully when producing words starting with a voiced plosive when there is a voiceless word competitor (because this reduces the chance that the voiced plosive will be mistakenly perceived as voiceless) than when there is no voiceless word competitor. The influence of the existence of a voiceless word competitor on prevoicing was therefore tested.

Method

Materials

Sixty-four items beginning with voiced plosives were selected. They were all monosyllabic. In the materials the following factors were varied: the place of articulation of the plosive (labial versus alveolar); the phoneme following the plosive (vowel versus consonant); the lexical status of the item (word versus non-word); and the competitor environment of the item, that is, whether changing the first voiced plosive into its voiceless counterpart resulted in a word or a nonword (competitor versus no competitor). The vowels that followed the plosives were: /a/, /ɑ/, /o/, /ɔ/, /i/, /ɪ/, /e/, /ɛ/, /œy/ and /eu/. The consonants that followed the plosives were: /l/, /r/, or /ʋ/. All four factors were fully crossed, resulting in 16 conditions. Each condition contained four items. Table 2-1 shows the full design and an example of each combination of factors. The full set of materials is listed in Appendix 2-A.

In addition to the 64 test items there were 456 fillers, resulting in a list of 520 items. The group of fillers contained both mono- and bisyllabic words and nonwords. The fillers were added to prevent the participants' attention from being drawn to the stimuli starting with voiced plosives. Some of the filler items served as test items for Experiment 2.2. Approximately one third of the items on the list started with a voiced plosive.

Participants

Participants were students from the MPI subject pool. There were five male and five female speakers. All of them were native speakers of Dutch and fluent readers. They were paid for their participation.

Recordings

Participants were seated in a sound-proof booth and were asked to read the items on the list out loud in front of a microphone, which was placed approximately 30 cm from the mouth. The items were presented without any context and participants were instructed to read the items one by one, separated by a pause, in a clear and natural way. If they made a mistake they could read the word again. Recordings were made onto digital audio tape (sampling rate of 48 kHz with 16-bit resolution). After applying an anti-alias filter, the utterances were redigitized at a sample rate of 16 kHz.

Table 2-1. *Full design of Experiment 2.1. Each combination of factors contained 4 items.*

Place of articulation	Lexical status	Following phoneme	Competitor environment	Item (competitor)	
Labial	Nonword	Vowel	No competitor	baag -	
			Competitor	bijn -	(pijn) *(pain)*
		Consonant	No competitor	bleep -	
			Competitor	bluim -	(pluim) *(feather)*
	Word	Vowel	No competitor	biels *(sleeper)*	
			Competitor	boot *(boat)*	(poot) *(paw)*
		Consonant	No competitor	brood *(bread)*	
			Competitor	bril *(glasses)*	(pril) *(young)*
Alveolar	Nonword	Vowel	No competitor	daaf -	
			Competitor	daart -	(taart) *(pie)*
		Consonant	No competitor	dwomp -	
			Competitor	draan -	(traan) *(tear)*
	Word	Vowel	No competitor	deur *(door)*	
			Competitor	duin *(dune)*	(tuin) *(garden)*
		Consonant	No competitor	dwars *(diagonally)*	
			Competitor	drol *(turd)*	(trol) *(troll)*

Measurements

For each token the duration of prevoicing was measured. The beginning of the prevoicing was defined as the point in time at which evidence of vocal fold vibration could be detected. Any clearly visually detectable period, no matter how small in amplitude, was accepted as part of voicing. The end of the prevoicing was defined as the point in time at which the noise of the release burst started, visible as a sudden peak in the waveform. Only when it was not completely clear where the prevoicing or the plosive release started, a wide-band spectrogram was used to locate the point in time where there was a sudden presence of aperiodic wide-band energy. We found three different prevoicing patterns: no prevoicing; voicing interrupted by the plosive release; or voicing continued during the release, in which case the release was visible as a short-term turbulent structure of low energy on top of the voicing pulses. Below we do not distinguish between the two latter patterns.

Results and Discussion

Figure 2-1 shows the percentage of prevoiced tokens per speaker. There was considerable variation between subjects: some speakers produced prevoicing at the beginning of each voiced plosive, while other speakers only did so for some of the items. One speaker produced only 38% of the items with prevoicing. Overall 75% of the tokens were produced with prevoicing.

First, the influence of the speakers' sex on the proportion of prevoiced tokens (see Figure 2-1) and on the duration of prevoicing was examined.

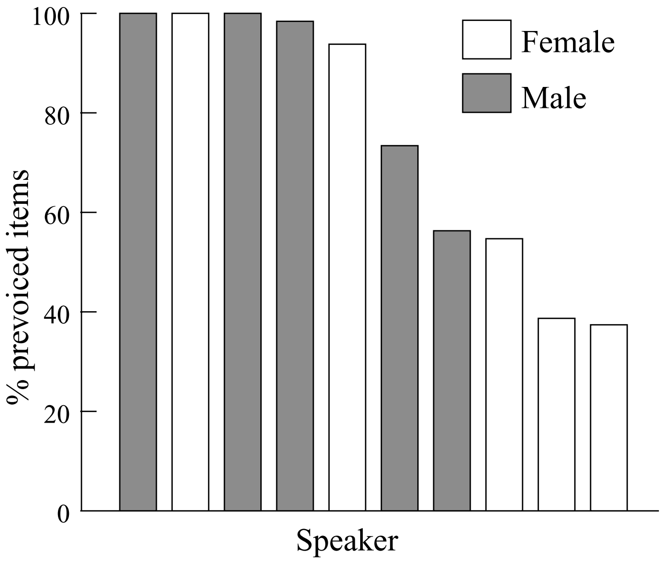

Figure 2-1. *Percentage of prevoiced items plotted separately for each speaker in rank order*

To investigate the influence of this factor on the proportion of prevoiced tokens, a logistic regression (LR) analysis was performed with prevoicing (present or absent) as the dependent variable and sex (male or female) as factor. The LR model with sex as factor plus a constant yielded a deviance of $G^2 = 675$ (residual df = 635), which was a significant improvement over the model consisting of only a constant ($G^2 = 711$; residual df = 636). The coefficient for the sex of the speaker was significantly different from zero ($B = -1.2$, $p < .0001$). As predicted, male speakers produced more tokens with prevoicing than female speakers did (86% versus 65% respectively). To investigate the influence of sex on prevoicing duration a one-way analysis of variance (ANOVA) on the prevoicing duration of only the prevoiced tokens was performed. The difference in prevoicing duration between males (109 ms) and females (89 ms) was not significant.

Second, the influence of the four factors (place of articulation, lexical status, following phoneme and competitor environment) on the proportion of prevoicing and the prevoicing duration was examined. The mean percentages of prevoiced tokens calculated separately for each of the four factors are shown in column 3 of Table 2-2. As before, a logistic regression analysis with prevoicing (present or absent) as the dependent variable was performed. This time there were four independent variables: place of articulation (labial or alveolar), lexical status (word or nonword), following phoneme (vowel or consonant) and competitor environment (no competitor or competitor).

Table 2-2. *Percentage of prevoiced tokens for all ten speakers and mean prevoicing duration (in milliseconds) of the prevoiced tokens of the five most frequent prevoicers*

Factor	Level of factor	% Prevoiced tokens	Prevoicing duration mean	(sd)
Place of articulation	Labial	78.9	112.9	(32.2)
	Alveolar	71.8	104.1	(23.0)
Following phoneme	Vowel	85.5	117.5	(29.2)
	Consonant	65.3	99.5	(24.2)
Lexical status	Nonword	76.8	109.7	(26.4)
	Word	73.9	107.3	(30.1)
Competitor environment	No competitor	73.9	105.5	(28.0)
	Competitor	76.8	111.5	(28.3)

The LR model with these four factors plus constant (G^2 = 669; residual df = 632) was significantly better than the model with only a constant (G^2 = 711; residual df = 636). Of the four factors, only two were significant. These were place of articulation (B = .21, $p < .05$) and following phoneme (B = -.58, $p < .0001$). Labial plosives were more often produced with prevoicing than alveolar plosives, and plosives followed by a vowel were more often produced with prevoicing than plosives followed by a consonant. The two lexical factors (lexical status and the competitor environment of the carrier stimulus) did not have a significant effect on the presence or absence of prevoicing.

Subsequently, we focused on the tokens with prevoicing to find out which of the four factors had an influence on the duration of prevoicing of these tokens. Since some speakers produced too few tokens with prevoicing to conduct a four-way repeated measures analysis of variance, we selected the five strongest prevoicers, that is, the speakers who produced more than 90% of the items with prevoicing. Only tokens produced with prevoicing were included in the analysis. Column 4 of Table 2-2 shows the mean duration of the prevoicing for the four factors separately, collapsed over these five frequent prevoicers. Following phoneme was the only factor showing a significant main effect: F1(1,4) = 63.6, $p < .001$; F2(1,4) = 24.8, $p < .001$. The duration of the prevoicing was longer for plosives followed by a vowel (118 ms) than for plosives followed by a consonant (99 ms). The effect of the place of articulation was significant in the items analysis: F2(1,4) = 6.01, $p < .05$, but did not reach significance in the subjects analysis: F1(1,4) = 4.95, $p = .09$. There were no significant effects of the lexical factors. There was, however, a significant three-way interaction of following phoneme, lexical status and word competitor: F1(1,4) = 16.2, $p < .05$; F2(1,4) = 7.06, $p < .05$, but a post hoc Tukey honestly significant test showed that there were no significant pairwise differences in the items analysis.

In summary, we found much variation in prevoicing of initial Dutch voiced plosives among speakers. Overall, about 75% of the tokens were prevoiced. Some speakers always produced prevoicing, but others did so in less than half of the cases. Female speakers produced less tokens with prevoicing than male speakers. There was however no sex difference in prevoicing duration. Of the four explored factors, only the two segmental factors had a significant effect on the percentage of prevoiced tokens. When the place of articulation of the initial voiced plosive was labial, tokens were more often prevoiced than when the place of articulation was alveolar. Furthermore, prevoicing was omitted more often when the following phoneme was a consonant than when it was a vowel. The following phoneme also had an effect on the duration of the prevoicing: the duration was shorter for plosives followed by a consonant than for plosives followed by a vowel. There was no effect of either the lexical status of the stimulus or the lexical competitor environment on prevoicing production.

Our study confirms the finding by Lisker and Abramson (1964) that some speakers realize all voiced plosives with prevoicing. It also shows, however, that other speakers do not always produce prevoicing. Overall, 25% of all voiced plosives were produced without prevoicing.

We predicted that both the sex of the speaker and the place of articulation of the plosive would affect prevoicing production. These predictions were confirmed. Male speakers produced prevoicing more often than females did and labial plosives were more often produced with prevoicing than alveolar plosives. The effect of the sex of the speaker is probably due to differences in the size of the vocal tract between men and women. Men tend to have larger vocal tracts than women (e.g., Stevens, 1998) and therefore the supraglottal pressure rises less quickly in the former. This makes it easier to produce prevoicing. The difference in the duration of prevoicing which was present was, however, not significant.

The effect of the place of articulation on the occurrence of prevoicing can be explained by differences in the size of the surface of the vocal tract walls which can participate in the passive expansion. Since labial plosives are produced more anteriorly than alveolar plosives, the surface of tissue which can be pushed outward as a result of the raised oral pressure is larger for labials than for alveolars. In line with this, labials were more often produced with prevoicing than alveolars. There was no effect on the duration of the prevoicing.

In addition to these two predicted effects there was also an effect of the following phoneme; prevoicing was more often produced and, if present, longer when the plosive was followed by a vowel than when it was followed by a consonant. The group of following consonants consisted of three different phonemes: /r/ after /b/ and /d/, /l/ after /b/ and /ʋ/ after /d/. Although in some cases, the following consonant may result in a smaller size of the oral cavity, for example when the plosive /b/ is followed by /l/ in comparison to when it is followed by an /a:/, this explanation would not hold for all consonants and vowels which were used in this study. Furthermore, when we studied the occurrence and duration of prevoicing in the vowels separately, no effect of vowel height was found. Based on the above, it is very unlikely that the observed difference in prevoicing between plosives followed by a vowel and plosives followed by a consonant is only caused by a difference in the volume of the oral cavity. We therefore propose that the degree to which the vocal tract can be expanded (passively or actively) plays a role in these findings. Articulatory measurements should be obtained in order to find a detailed explanation for this effect.

Finally, the two lexical factors appeared to have no influence on prevoicing production. The finding that nonwords and words did not differ in the production of prevoicing show that nonwords are not hyperarticulated in the sense of more reliable prevoicing. Furthermore, the absence of an effect of the competitor environment indicates that it is not the case that listeners articulate more carefully to avoid activation of a voiceless word competitor.

Given that a quarter of the voiced plosives were produced without prevoicing, the question emerges whether these plosives are still perceived as voiced. Is the production of prevoicing essential for the plosives to be perceived as voiced, or are other acoustic cues present and strong enough to evoke a voiced percept? To answer these questions, first a detailed acoustic analysis of the productions of voiced and voiceless plosives was conducted in Experiment 2.2. Several potential acoustic cues were measured and analyzed. A classification tree analysis was performed to investigate which of the measured cues would be the most reliable for categorization of the voiced-voiceless distinction. Experiment 2.3 was designed to find out whether listeners identified the produced tokens as voiced or voiceless and which of the measured cues influenced identification most strongly.

EXPERIMENT 2.2

Based on the study by Slis and Cohen (1969), and the information on the voicing distinction in other languages (mainly English), the following six measurements were selected for the purpose of the present study: duration of prevoicing, duration of the burst, power of the burst, spectral centre of gravity of the burst, F_0 immediately after burst offset, and F_0 movement into the vowel.

Method

Materials

From the complete collection of 520 tokens produced by 10 different speakers (Experiment 2.1) 48 item pairs were selected. These pairs differed only in the voicing of the initial plosive, in order to obtain the same variation in segmental context in both groups (voiced and voiceless). Note, however, that these items were not produced as pairs, but as single items in random order among many fillers. Half of the pairs started with labial plosives (/b/ or /p/) and the other half started with alveolar plosives (/d/ or /t/). Half of the pairs started with a consonant cluster (/b/ or /p/ followed by an /r/ or /l/, and /d/ or /t/ followed by /r/ or /ʋ/) and the other half of the pairs started with a plosive followed by a vowel. One third of the pairs were nonword-word pairs, i.e., the voiced counterpart of the pair was a nonword and the voiceless counterpart a word, for example *bluim* -

pluim (plume); one third of the pairs were word-nonword pairs, for example *braam* (blackberry) - *praam*; one third of the pairs were word-word pairs, for example *baars* (perch) - *paars* (purple). The complete set of items is given in Appendix 2-B.

Measurement procedures

For each item produced by each speaker the six cues that were expected to signal the voiced-voiceless distinction were measured. Below, first the relevant references in Dutch are given, followed by a description of how the measurements were performed for each cue.

Duration of prevoicing

As already mentioned, Lisker and Abramson (1964) found that all tokens of voiced Dutch plosives produced by one speaker were prevoiced. Experiment 2.1, however, showed that only 75% of the tokens starting with a voiced plosive were produced with prevoicing. In a perceptual experiment using synthetic CV stimuli that varied only in VOT, Slis and Cohen (1969) found that voiced judgements correlated with a voice lead and voiceless ones with a voice lag. The methods of measuring the duration of the prevoicing were the same as in Experiment 2.1 (see Method of Experiment 2.1). Based on the results of Experiment 2.1 we expected to find prevoicing in approximately 75% of the voiced plosives. No prevoicing was expected in the productions of the voiceless plosives.

Duration of the burst

Slis and Cohen (1969) reported that the noise burst duration of Dutch plosives was on average 15 ms shorter for voiced plosives than for voiceless plosives. The difference in burst duration may be explained by the spatially more extended contact at constriction for voiceless plosives in comparison to voiced plosives (e.g., Cho & Ladefoged, 1999; Yoshioka, Murase & Uematsu, 1996). Ernestus (2000) measured burst durations of 649 Dutch plosives in medial position and showed that expert listeners tended to classify plosives with short bursts durations as voiced, and plosives with long burst durations as voiceless.

The onset of the burst was defined as the point in time at which the closure was released (see Experiment 2.1). The definition for the offset of the burst varied with the following phoneme. When the following phoneme was a vowel or an /l/, the offset of the burst was defined as the point at which higher formants were first visible in the spectrogram. When the following consonant was a /ʋ/ the offset of the burst was defined at the point where the spectrogram showed a sudden change in spectral composition of the noise. When the plosive was followed by an /r/ the labeling of the end of the burst depended on the way the /r/

was produced. In the cases where it was produced as an retroflex approximant [ɻ] the onset of higher formants served as an indication of the offset of the burst. In the cases were it was produced as an uvular trill [R] or alveolar trill [r] or as an uvular fricative [ʁ], the change in the structure of the noise served as indication for the burst offset. In many of these latter cases, the trill or frication was preceded by a short schwa. The moment at which the higher formants of the schwa were visible in the spectrogram were then taken as the offset of the burst.

The burst included the following two acoustic events, as described by Stevens (1993): a brief transient as the air that has been compressed in the vocal tract discharges through the opening constriction, followed by frication noise, which is caused by rapid airflow through the constriction. Dutch voiceless plosives have little or no aspiration. In the cases where there was aspiration, we included the aspiration in the burst. The duration of the burst was expected to be longer for voiceless plosives than for voiced plosives. Note that for voiceless plosives, the duration of the burst reflects the positive VOT.

Power of the burst above 500 Hz

Slis and Cohen (1969) found that the amplitude of the voiceless noise burst was about 50% higher than the amplitude of the voiced noise burst. Possible causes for this difference mentioned in the literature are higher oral pressure behind the constriction and/or spatially more extended closure for voiceless plosives (e.g., Yoshioka, Murase & Uematsu, 1996).

The burst was first high-pass filtered at a cutoff frequency of 500 Hz. Then the spectral power was calculated by taking the logarithm of the mean sum of squares of all sample points. Energy under 500 Hz was filtered out to exclude the energy generated by any vocal fold vibration during or immediately after the release of the closure. The spectral power was expected to be higher for bursts of voiceless plosives than for voiced plosives.

Spectral centre of gravity of the burst

The spectral centre of gravity (SCG), or first spectral moment, has been used to describe the difference in place of articulation of fricatives and plosives (e.g., Forrest, Weismer, Milenkovic & Gougall, 1988), since it is used as an acoustic measure for the size of the front cavity; the smaller the front cavity, the higher the SCG. We propose that the SCG is also an appropriate measure for the difference in between voiced and voiceless plosives for several reasons. First, we predict that the SCG is influenced by the presence of voicing in the burst. When there is voicing, there is more energy in the lower frequencies which will shift the gravity to lower frequencies compared to when there is no voicing. Therefore we expected the SCG to be lower in voiced plosives than in voiceless plosives. Second, Cho, Jun and Ladefoged (2002) remark in a study of alveolar fricatives that the SCG might also reflect the velocity of the jet of air; a higher subglottal pressure results in a jet of air with a higher velocity, which will result in a higher SCG. Voiceless plosives are expected to be produced with a higher velocity of the air jet than voiced plosives, which would result in higher SCG for the voiceless plosives than for the voiced plosives. Finally, our intuition (as native speakers of Dutch) is that the place of articulation of /d/ is slightly different from that of /t/. The voiceless counterpart seems to be produced more frontally than the voiced one. This suggests that the front cavity is smaller for /t/ than for /d/, resulting in higher SCG for the voiceless alveolar plosive. If this is indeed true, we should find that the difference in SCG between voiced and voiceless plosives is larger in the case of alveolar plosives than in the case of labial plosives.

To calculate the SCG, the burst was first filtered into 32 frequency bins with widths of 250 Hz, except for the first bin which was high-pass filtered at a cut off of 50 Hz (to remove any spurious low frequency components) resulting in a range from 50 to 250 Hz. For each filter the power of the filtered signal was calculated. Next, the 32 centre frequencies were multiplied by the corresponding powers, summed together and divided by the sum of the powers, resulting in the SCG.

Absolute F_0 and F_0 difference

Many studies have reported a higher fundamental frequency (F_0) of the vowel adjacent to a voiceless plosive than of the vowel adjacent to a voiced plosive in English (House & Fairbanks, 1953; Kingston & Diehl, 1994; Lehiste & Peterson, 1961; Mohr, 1971; Löfqvist, 1975; Umeda, 1981). For Dutch, Slis and Cohen (1969) reported a difference of 6 Hz between the top frequency after voiceless consonants and the top frequency after voiced consonants. A possible cause for the difference in F_0 is the lowering of the larynx during the production of voiced plosives in order to obtain sufficient transglottal pressure to produce vocal fold vibration. Lowering of the larynx can cause a downward tilt of the cricoid cartilage, which causes a shortening and hence slackening and thickening of the vocal folds (Honda, Hirai & Kusakawa, 1993), resulting in a lower F_0. This explanation suggests that the F0 pattern is directly related to the moment at which the vocal fold vibration starts. Ohde (1984) reported, however, that in English a high F_0 was found for both voiceless aspirated (in initial position) and unaspirated plosives (after an /s/ in initial position), although the VOTs of those groups were very different. These data suggest that F_0 differences are a product of articulations that are controlled independently of the timing of the glottal articulations to produce voicing (see also Kingston & Diehl, 1994). In sum, the relationship between voicing and F_0 patterns remain contoversial.

Haggard, Ambler and Callow (1970) demonstrated that stimuli were consistently perceived as /b/ when synthesized with a low-rising F_0 contour, but as /p/ with a high-falling contour. Further perception data by Haggard, Summerfield and Roberts (1981) suggest that the actual cue is the onset frequency rather than the F_0 movement into the vowel. Many other studies have examined the influence of F_0 differences on the perception of the voicing distinction in plosives (e.g., Ohde, 1984; Kohler, 1985; Whalen, Abramson, Lisker & Mody, 1993), but the underlying perceptual mechanisms remain largely unknown. Both the absolute F_0 value immediately after the plosive and the F_0 movement into the vowel were therefore included as potential cues in the present study.

F_0 was estimated for each frame of 10 ms of the vowel (or consonant plus vowel) that followed the initial plosive, using an algorithm called RAPT (Talkin, 1995), which estimates the fundamental frequency from the normalized cross correlation function using dynamic programming. Subsequently, the mean F_0 was calculated for the first two voiced frames, resulting in a measure for the absolute F_0 immediately after burst offset. This absolute F_0 was expected to be lower for voiced plosives than for voiceless plosives.

To obtain a measure of F_0 change, the F_0 immediately after burst offset (see above) was subtracted from the F_0 in the middle of the vowel in the cases were the plosive was followed by a vowel, or from the F_0 from the middle of the consonant plus the following vowel in the cases were the plosive was followed by a consonant. Thus, a positive F_0 difference corresponded to a rising F_0 pattern and a negative F_0 difference corresponded to a falling F_0 pattern. The F_0 in the middle was defined as the mean F_0 for the middle two or three frames (depending on whether the total number of frames was even or odd). The two or three middle frames were not allowed to overlap with the two frames used to estimate the F_0 immediately after burst offset. Note that therefore the vowel (or second consonant plus vowel) was required to have a duration of at least 60 ms (6 frames). For the tokens which did not meet this constraint, no F_0 difference was calculated.

RESULTS AND DISCUSSION

The measures in the current experiment were selected in order to describe the voiced-voiceless distinction. However, on the basis of previous literature, we expected that most of these measures would also vary with place of articulation. Therefore, in addition to the voicing category of the plosive, the place of articulation of the plosive was included in the data analyses. In some cases not all measurements could be obtained, for example when the recording was affected by any external noise, or when it was not clear where a particular segment started or ended.

The distributions of each measure are plotted separately for /b/ versus /p/ and /d/ versus /t/ in Figures 2-2 to 2-7. To find out whether these measures were good correlates of the voicing distinction, a multivariate analysis of variance was conducted with the six measures as dependent variables and voicing category (voiced versus voiceless) and place of articulation (labial versus alveolar) as factors. Table 2-3 indicates the significant main effects and interactions. These effects will be discussed for each measure separately.

The duration of prevoicing was longer for voiced than for voiceless items. In fact, as expected, none of the voiceless items were produced with prevoicing. Therefore, only the histograms for the voiced plosives were plotted (see Figure 2-2). Overall, 76% of the voiced tokens were produced with prevoicing.

The duration of the burst was longer for voiceless plosives than for voiced plosives (Figure 2-3). The difference between the burst duration of voiced and voiceless plosives was approximately 10 ms. For the voiceless plosives the duration of the burst reflects the positive VOT. The mean VOT values for the voiceless plosives were almost twice as high as the VOT values for Dutch reported by Lisker and Abramson (1964), but were similar to the VOT values reported by Flege and Eefting (1987). Place of articulation also had an effect on burst duration: bursts were longer for alveolars than for labials. This is in line with findings in the literature for Dutch initial plosives (Smits, 1995).

The power of the burst above 500 Hz was higher for voiceless plosives than for voiced plosives (Figure 2-4). The difference between voiced and voiceless plosives was about 3 dB. The power of the burst was also higher for alveolar plosives than for labial plosives.

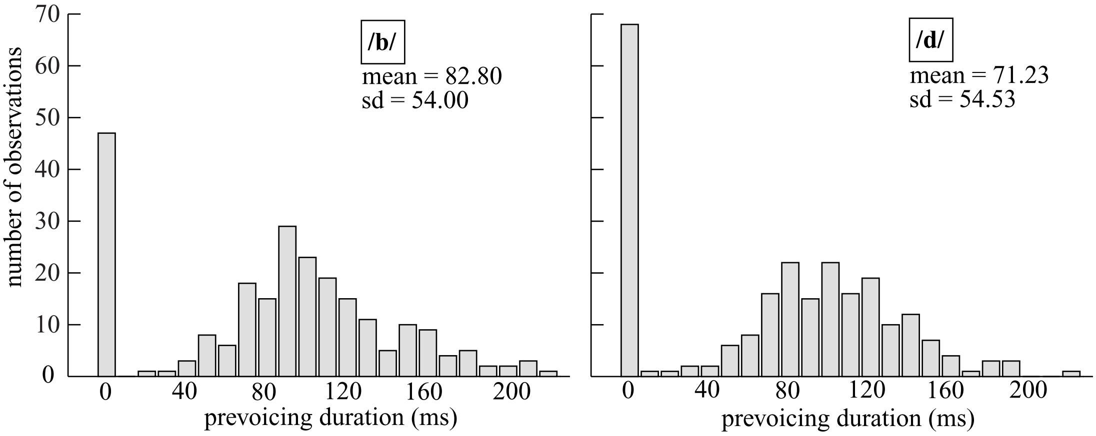

Figure 2-2. *Histograms of the prevoicing as produced by 10 different speakers, plotted separately for place of articulation (left versus right). Only the voiced categories are plotted; all tokens of the voiceless category were produced without prevoicing. The numbers on the x-axis represent the upper limits of each bin.*

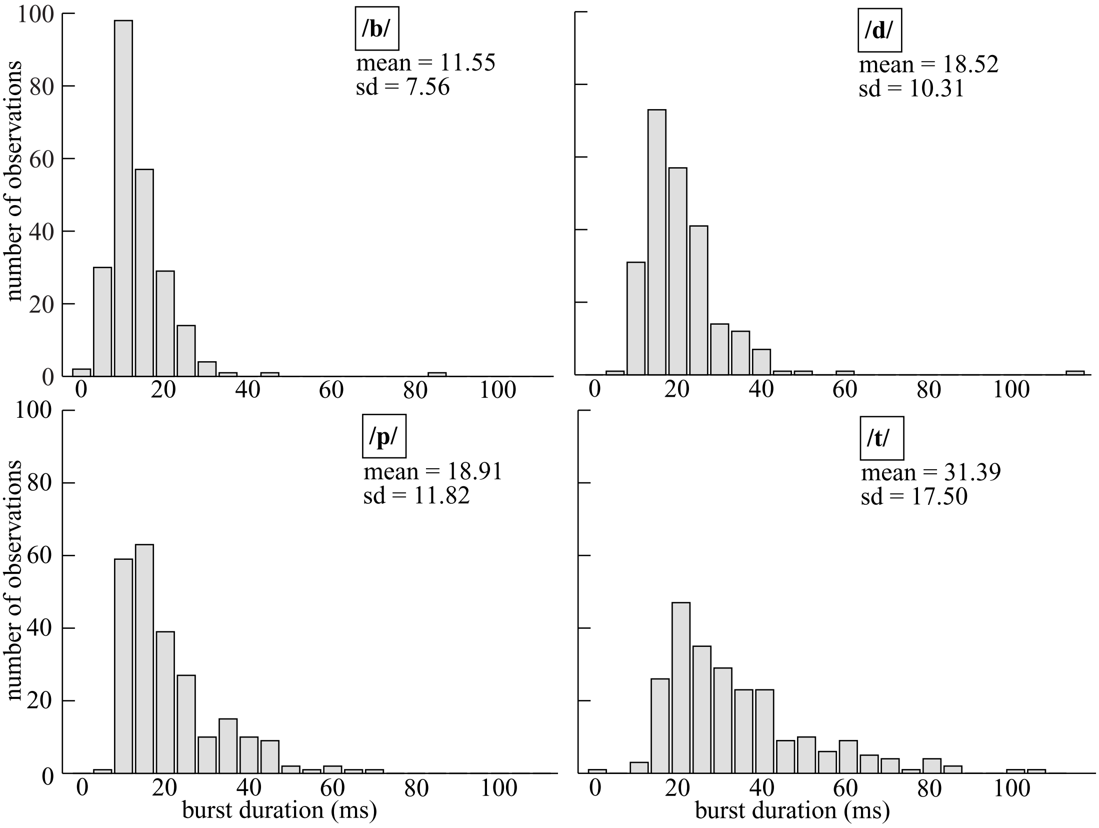

Figure 2-3. *Histograms of the burst durations as produced by 10 different speakers, plotted separately for place of articulation (left versus right) and voicing category (top versus bottom). The numbers on the x-axis represent the upper limits of each bin.*

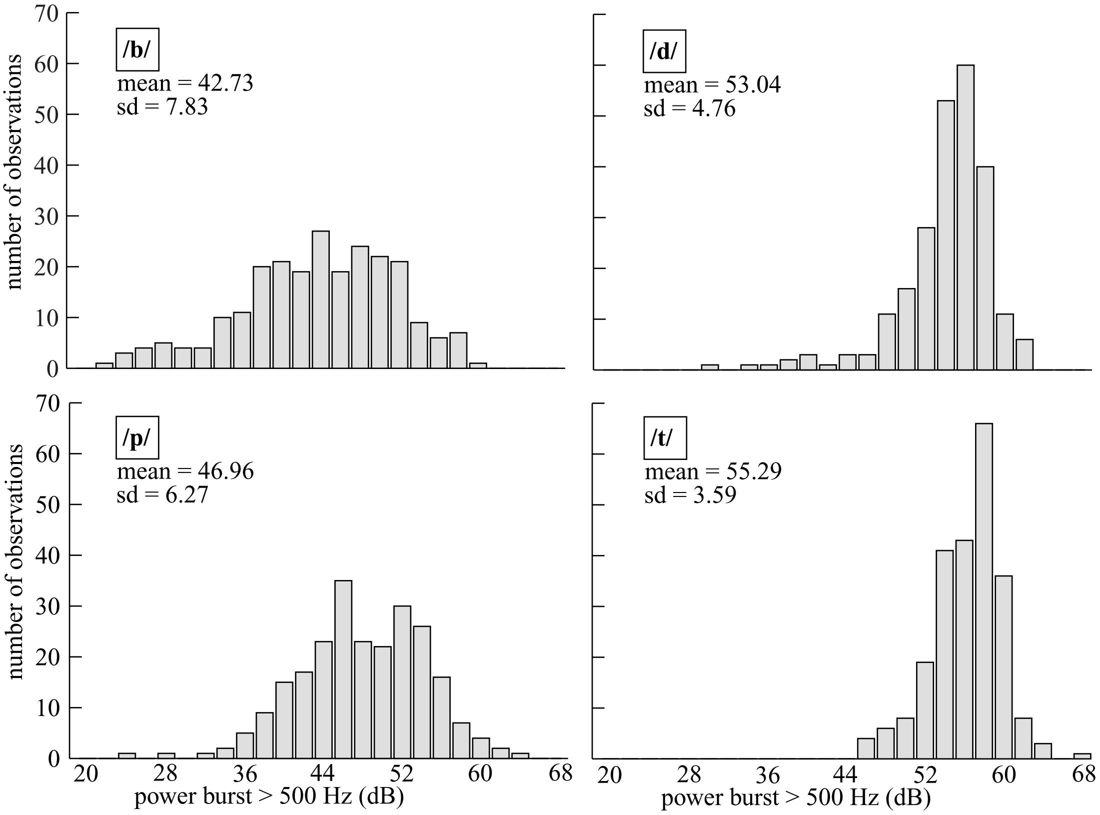

Figure 2-4. *Histograms of the power in the burst as produced by 10 different speakers, plotted separately for place of articulation (left versus right) and voicing category (top versus bottom). The numbers on the x-axis represent the upper limits of each bin.*

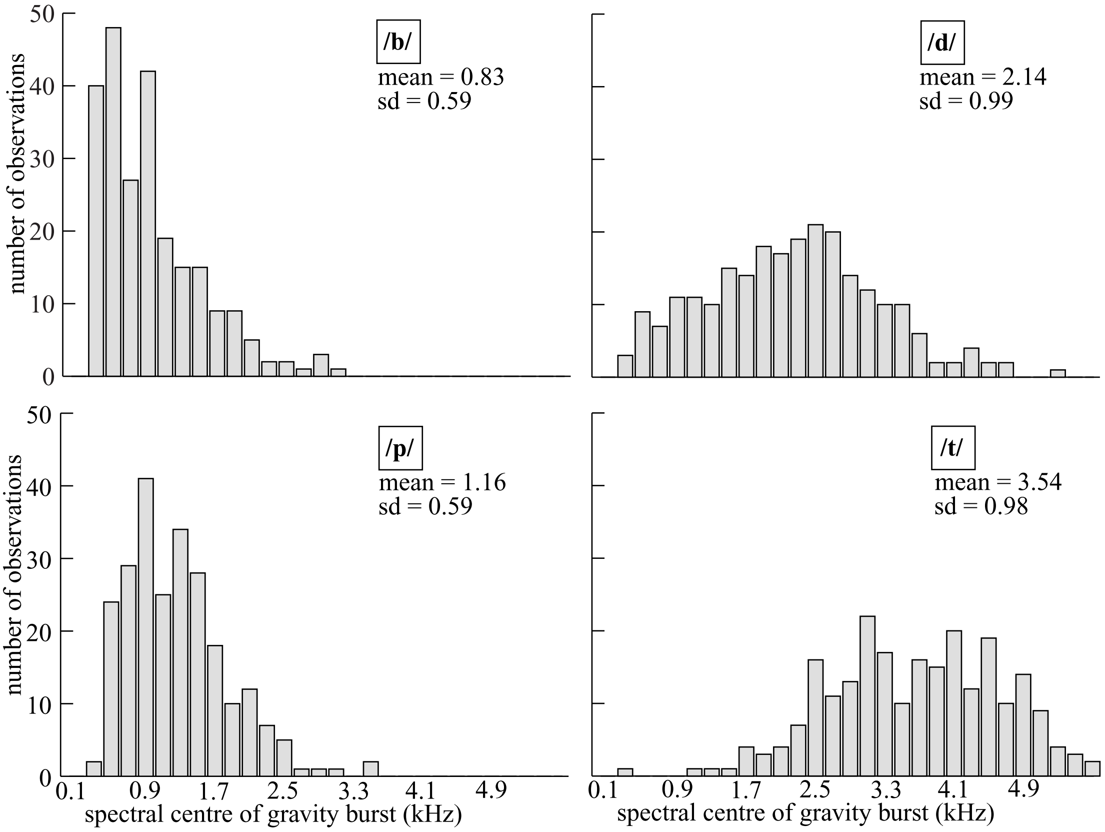

Figure 2-5. *Histograms of the spectral centre of gravity of the burst as produced by 10 different speakers, plotted separately for place of articulation (left versus right) and voicing category (top versus bottom). The numbers on the x-axis represent the upper limits of each bin.*

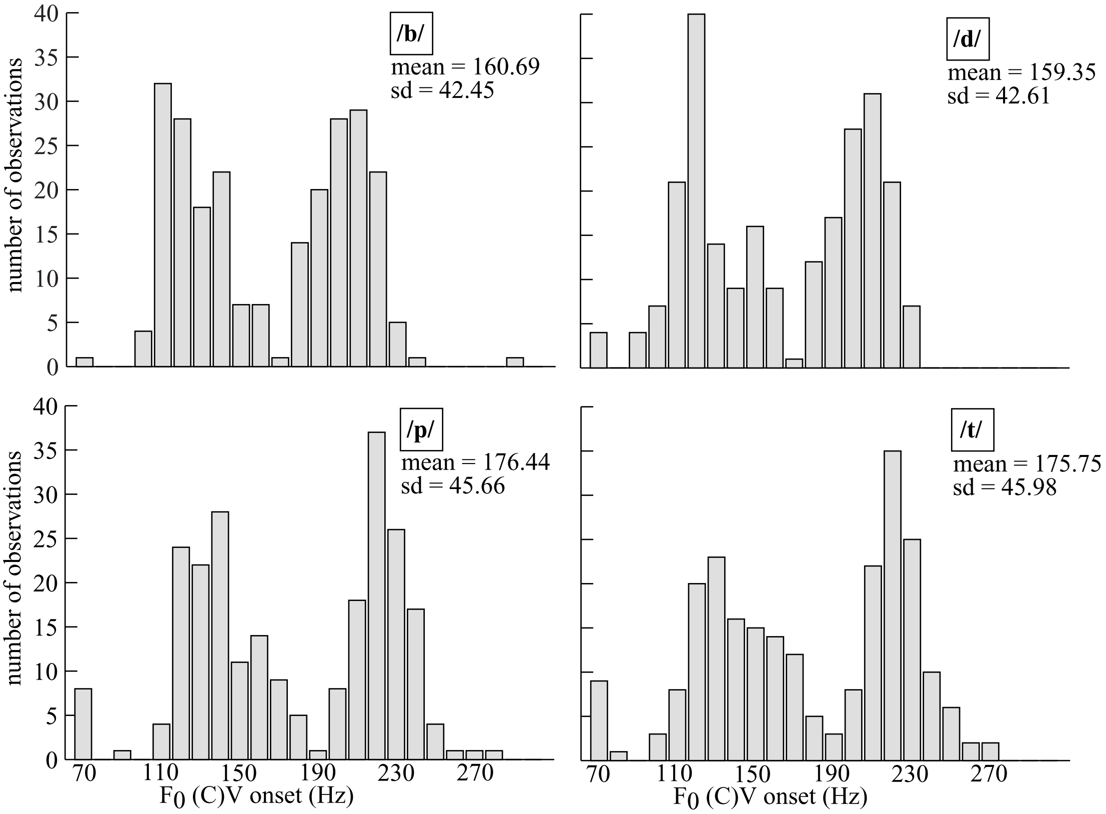

Figure 2-6. *Histograms of the F_0 at (C)V onset as produced by 10 different speakers, plotted separately for place of articulation (left versus right) and voicing category (top versus bottom). The numbers on the x-axis represent the upper limits of each bin.*

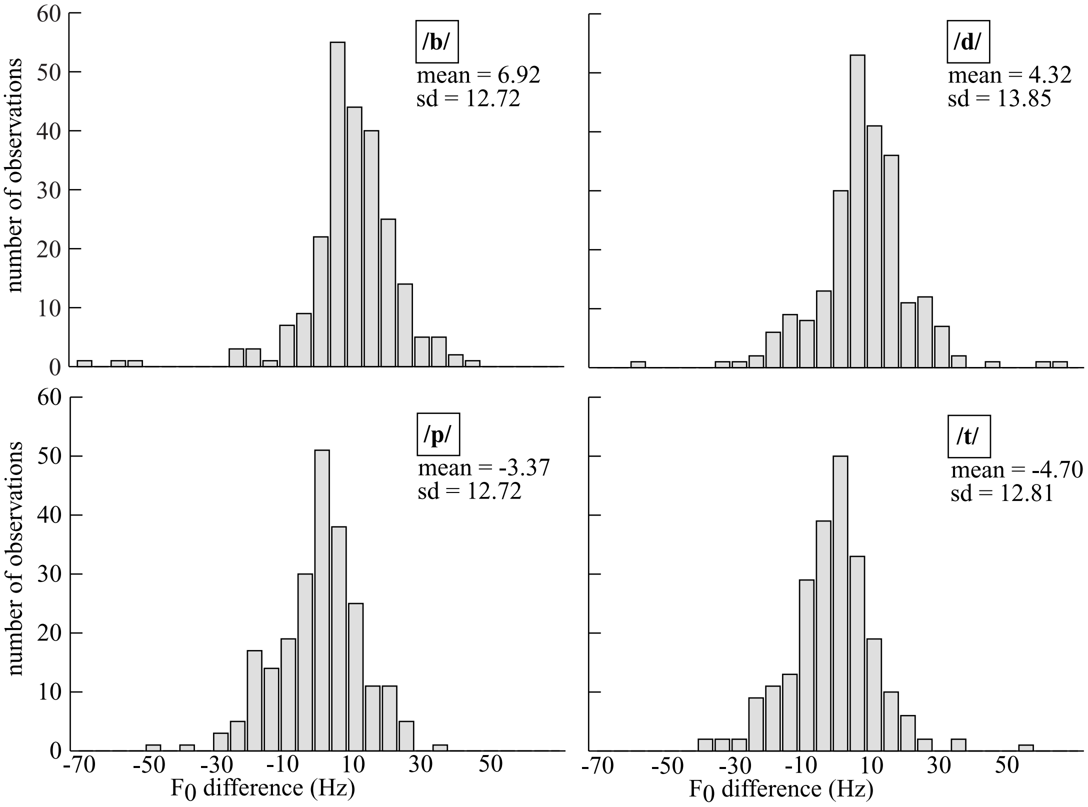

Figure 2-7. *Histograms of the* F_0 *difference as produced by 10 different speakers, plotted separately for place of articulation (left versus right) and voicing category (top versus bottom). The numbers on the x-axis represent the upper limits of each bin.*

Table 2-3. *Significant effects of the multivariate ANOVAs (subjects and items analyses) for the six measures.*

	Voicing	Place of articulation	Voicing x Place of articulation
Prevoicing duration (ms)	F1(1,36) = 92.53 p < .0001 F2(1,92) = 623.48 p < .0001	F1(1,36) = .54 not significant F2(1,92) = 3.42 not significant	F1(1,36) = .54 not significant F2(1,92) = 3.42 not significant
Burst duration (ms)	F1(1,36) = 23.63 p < .001 F2(1,92) = 47.90 p < .0001	F1(1,36) = 21.83 p < .0001 F2(1,92) = 44.25 p < .0001	F1(1,36) = 1.72 not significant F2(1,92) = 3.57 not significant
Power of burst (dB)	F1(1,36) = 11.85 p < .0001 F2(1,92) = 63.20 p < .001	F1(1,36) = 97.57 p < .0001 F2(1,92) = 523.66 p < .0001	F1(1,36) = 1.08 not significant F2(1,92) = 5.95 p < .05
SCG (kHz)	F1(1,36) = 52.09 p < .0001 F2(1,92) = 155.91 p < .0001	F1(1,36) = 238.07 p < .0001 F2(1,92) = 713.77 p < .0001	F1(1,36) = 20.35 p < .0001 F2(1,92) = 60.28 p < .0001
F_0 at burst offset (Hz)	F1(1,36) = 1.42 not significant F2(1,92) = 158.71 p < .0001	F1(1,36) = .00 not significant F2(1,92) = .66 not significant	F1(1,36) = .00 not significant F2(1,92) = .08 not significant
F_0 difference (Hz)	F1(1,36) = 16.57 p < .0001 F2(1,92) = 110.38 p < .0001	F1(1,36) = .73 not significant F2(1,92) = 4.39 p < .05	F1(1,36) = .06 not significant F2(1,92) = .52 not significant

Figure 2-5 clearly shows that the distributions of the SCG are very different depending on the place of articulation. Mean SCG was higher for alveolars (2.8 kHz) than for labials (1.0 kHz), as has been described in the literature (e.g., Forrest, Weismer, Milenkovic & Dougall, 1988). As predicted, the SCG appeared also to be influenced by the voicing category of the plosive: it was higher for voiceless plosives than for voiced plosives.

In addition to the two main effects, there was an interaction between voicing and place of articulation, which shows that the difference in the SCG between voiced and voiceless was considerably larger for alveolars (1.40 kHz) than for labials (0.32 kHz). At first sight it may seem that this interaction can be explained by the fact that labial plosives are more often produced with prevoicing than alveolar plosives. One might predict that when more tokens are produced with prevoicing also more of the bursts would be voiced, which would result in lower SCG values for prevoiced tokens. However, the asymmetry in the presence of prevoicing in labials and alveoars would result in the opposite pattern, namely in a larger difference in SCG between voiced and voiceless labials than between voiced and voiceless alveolars. We also discussed how the SCG is related to the front cavity. The present observed interaction is in agreement with our intuition that voiced alveolar plosives are produced slightly more posteriorly than the voiceless counterpart. Of course, this hypothesis will have to be correlated with independent articulatory evidence, but the observed difference is in line with the predictions. The front cavity would be larger for /d/ than for /t/, resulting in a larger SCG difference between voiced and voiceless plosives in the case of alveolar plosives than in the case of labial plosives.

One possibility is that this difference in the place of articulation, and thus the size of the front cavity, is a by-product of the downwards displacement of the tongue body for the production of a voiced plosive (Svirsky et al. 1997). This downward displacement will enlarge the vocal tract, such that the air pressure rises less quickly and makes it therefore easier to produce prevoicing (Westbury, 1983). The downward displacement of the tongue body would also take place during the production of labial voiced plosives. In the cases of labials, however, displacement of the tongue does not affect the place of articulation, which is realized with the lips. Therefore, the size of the oral cavity will differ between voiced and voiceless labial plosives, but not the size of the frontal cavity. Another possible alternative explanation is that the difference in place of articulation may be the product of a planned enhancing strategy. Backing the place of articulation for the /d/ will lower the SCG, thus mimicking the effect of prevoicing (Keating, personal communication).

Figure 2-6 shows that the distribution of absolute F_0 immediately after burst offset is bimodal. This is caused by the difference in F_0 between male and female speakers. Although the difference between the mean F_0 of voiced and voiceless plosives was in the same range of 10 to 15% found by Ohde (1984) and by Kingston and Diehl (1994), this difference was not significant.

The mean F_0 difference (Figure 2-7) was positive for tokens starting with voiced plosives, consistent with a rising F_0, while it was negative for tokens starting with a voiceless plosive, consistent with a falling F_0. The difference between these means was significant. There also was an effect of place of articulation on the F_0 difference: the F_0 difference was larger for labials than for alveolars (1.84 versus -0.l3).

Taken together, the results show that the means for all tested acoustic properties showed the predicted patterns and that all, except for the F_0 at the offset of the burst, differed significantly between voiced and voiceless plosives. Furthermore, most of the measures differed between labial and alveolar plosives. For the SCG an interaction was found between voicing category and place of articulation. This suggests that there is a difference in the place of articulation between /d/ and /t/, but not between /b/ and /p/.

The acoustic analyses reveal that there are several acoustic properties which correlate with the voiced-voiceless distinction in Dutch plosives. The analyses do not show, however, which of these acoustic properties are most useful for correct recognition of the voicing feature. One would predict that listeners' phoneme identification would be influenced most by the cues which lead to the highest recognition scores. The obvious analyses to examine the relative strengths of the various acoustic properties for recognition are linear discriminant analysis or logistic regression analysis. These analyses are inappropriate for our data set, however, since most of the predictor variables were highly skewed or multi-modal. Moreover, we wanted to add a categorical predictor, namely whether the following phoneme was a vowel or a consonant. The suitable analysis for this type of data is a classification and regression tree (CART) analysis (Breiman, Friedman, Olshen & Stone, 1984). CART is a non-parametric analysis, i.e., no assumptions are made regarding the underlying distributions of the predictor variables. Furthermore, tree-based models are better than the traditional methods in managing complex interactions that may exist in the data .

Tree-based models operate by recursively partitioning a dataset in two (i.e., a binary split). Each split is based on the value of a single predictor variable. The choice of the predictor variable and its value for each split is based on an exhaustive search of all possible divisions of the data. The aim of each split is either to maximize the homogeneity of the groups in the case of nominal or ordinal responses (classification) or to best separate low and high values in the case of continuous response variables (regression). The algorithm continues splitting the subsets of data (the nodes) until they are maximally homogeneous or contain too few observations. Finally, the constructed tree is pruned using cross-validation, that is, the tree is simplified without sacrificing goodness-of-fit.

In the CART analysis we determined which of the acoustic correlates best predicted membership of the classes of voiced or voiceless plosives. The response variable that the analysis attempted to predict was the intention of the speaker to produce a voiced or voiceless plosive. Since the response variable was categorical, a classification tree analysis was conducted. All six acoustic cues were used as predictor variables. In addition to these numerical predictors, one categorical predictor was added, namely whether the plosive was followed by a vowel or consonant. Since the acoustic properties of voiced and voiceless plosives differed by place of articulation, separate classification tree analysis were conducted for labial and alveolar plosives. All tokens for which one or more of the measurements could not be made were excluded from the analyses (14 bilabials and 18 alveolars)

.

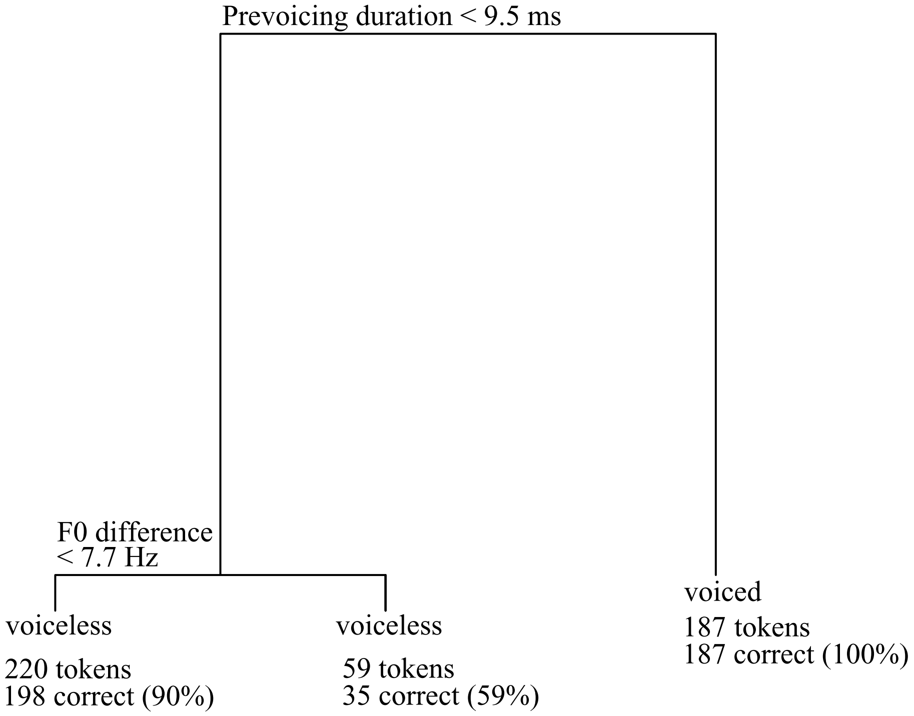

Figure 2-8. *CART analysis of the labial productions of the speakers. Tokens satisfying the rule printed at the top of each split followed the left branch. Final nodes are labeled according to the plurality rule. Below each label the total number of tokens falling in that final node and the number of correcly and incorrecly classified tokens is given.*

The resulting cost-complexity-pruned classification trees are shown in Figures 2-8 and 2-9. Each tree consists of a root node containing all tokens. This node is split based on a simple rule. Tokens satisfying the rule printed at the top of each split followed the left branch. The vertical length of each branch reflects the relevance of each factor, that is, the reduction in heterogeneity in each node. Each terminal node is labeled "voiced" or "voiceless" according to the plurality rule, i.e., according to the most represented class in that group of tokens. Below each label the number of tokens falling in that particular final node is given. The number of correctly and incorrectly classified tokens is printed underneath, which gives an indication of the goodness of fit of the tree.

The overall structure of the two trees is very similar. First the data is divided into two large groups on the basis of prevoicing duration, and then a small part of the data was further subdivided into two smaller groups. For the labial plosives, tokens produced with more than 9.5 ms of prevoicing followed the right branch and fell into a terminal node labeled "voiced". All tokens produced with more than 9.5 ms of prevoicing were intended by the speaker as being voiced plosives.

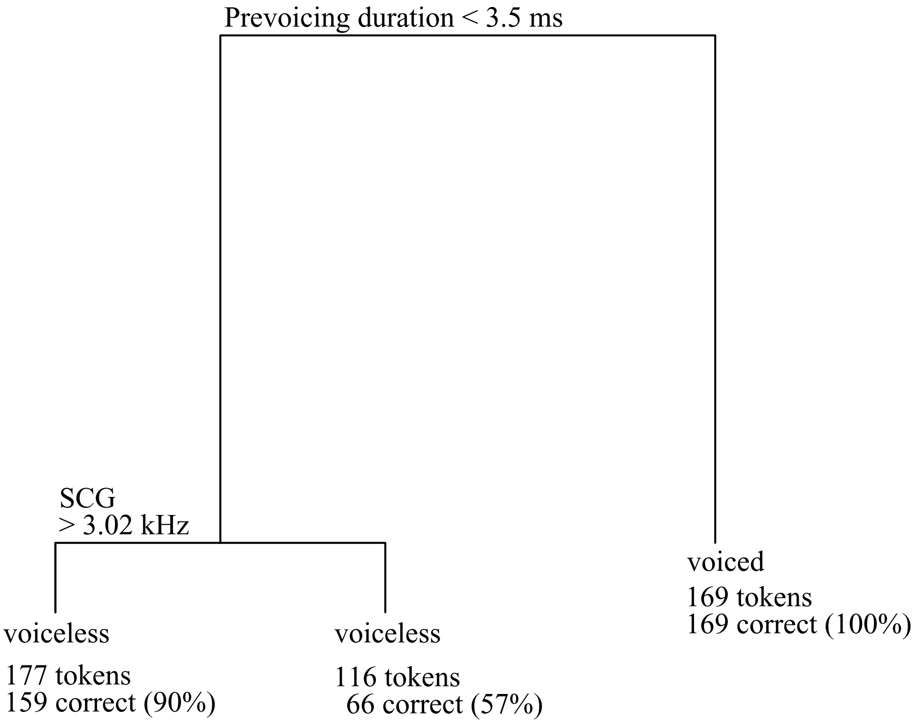

Figure 2-9. *CART analysis of the alveolar productions of the speakers. Tokens satisfying the rule printed at the top of each split followed the left branch. Final nodes are labeled according to the plurality rule. Below each label the total number of tokens falling in that final node and the number of correcly and incorrecly classified tokens is given.*

The same holds for the alveolar plosives which were first split based on a prevoicing duration of 3.5 ms. Since the cut-off values of 3.5 and 9.5 ms are barely enough to contain one period of voicing, we discuss the split according to whether the prevoicing duration was smaller or larger than 3.5 ms or 9.5 ms in terms of whether there was any prevoicing present or not.

The subset of labial productions without prevoicing was divided on the basis of the F_0 difference, while the subset of alveolar plosives was divided on the basis of the SCG. For the labials, 90% of the tokens without prevoicing and a small F_0 movement (< 7.7 Hz) was intended as being voiceless. Although the majority of the labials without prevoicing and a larger F_0 movement (that is, a clear rising F_0 pattern) was still intended as being voiceless, this proportion was now only 59%. In other words, on the basis of the acoustic measures, the CART analysis essentially distinguishes three groups of plosives: clearly voiced ones, clearly voiceless ones and an uncertain category of which a narrow majority is voiceless. For the alveolars, 90% of the tokens without prevoicing and a high SCG (> 3.02 kHz) were intended as being voiceless, while only 57% of the tokens without prevoicing and a lower SCG was intended as being voiceless. Again, the CART analysis thus finds three categories, one of which is uncertain. Overall 90% of the labial plosives and 85% of the alveolar plosives were correctly classified by the CART analysis.

Although the overall structure of both trees is similar, the proportion of the first main split and the second smaller split is different for the two places of articulation. For the labials, the vertical length of the branches of the prevoicing split is very long and the branches of the second split very small. The same is true for the alveolars, but in comparison to the labials the branches of the prevoicing split are somewhat smaller and the branches of the second split are somewhat longer.

In summary, our analysis of the production data indicated that there are several acoustic correlates to the voiced-voiceless distinction in Dutch. The classification tree analysis showed that the duration of prevoicing is by far the most reliable predictor of voicing, for both labial and alveolar plosives. All the tokens produced with prevoicing were intended as being voiced. For labials, the F_0 movement was the second most reliable predictor of voicing, while for alveolars this was the SCG. The fact that the CART analyses selected different acoustic cues to split the labial plosives without prevoicing and the alveolar plosives without prevoicing strengthens our claim that the acoustic realization of the voicing distinction differs for the two places of articulation. The strength of these acoustic cues (F_0 movement and SCG) is however small in comparison to the strength of prevoicing.

Experiment 2.2 examined which acoustic properties signal the voicing distinction and which of these acoustic correlates were the most reliable predictors of the voicing distinction as produced by speakers, but did not examine how the produced tokens were perceived by listeners. One would predict that there would be a good correspondence between the intended voicing category and the perceived voicing category, since the productions were produced naturally and clearly. Nevertheless, the results remain somewhat ambiguous. On the one hand, the presence or absence of prevoicing is the most reliable predictor of voicing, while on the other hand, a quarter of the voiced plosives were produced without prevoicing. The question therefore arises how voiced plosives that are produced without prevoicing are perceived by listeners. Is prevoicing indeed the strongest cue to the perception of the voicing distinction, as one would expect on the basis of the earlier acoustic analyses? It need not be true that listeners weight various cues in the same way as an automatic classifier would. What counts as important to the listener will of course depend on how the signal is processed by the peripheral auditory system and by the sensitivities of the speech perception system.

The purpose of Experiment 2.3 was therefore twofold. First we wanted to find out how the productions of Experiment 2.2 were perceived by listeners. In particular we were interested in the perceived voicing of the voiced plosives without prevoicing. Secondly, we wanted to explore which of the acoustic cues influenced the perception of the listener most strongly, and whether these cues corresponded to the cues which appeared to best describe the voicing distinction as produced by the speakers.

EXPERIMENT 2.3

Method

Materials

The materials for the perception experiment were based on the 48 voiced-voiceless pairs which were each produced by ten speakers and analyzed in Experiment 2.2. Only the initial portions of each token, up to the middle of the vowel, were presented, so as to prevent the listeners from using lexical information. To avoid creating abrupt amplitude changes, the offset of each fragment was ramped down to zero within a time-window of 10 ms.

Procedure

The fragments were presented binaurally over headphones at a comfortable listening level. The materials were blocked by place of articulation and speaker. This resulted in 20 blocks of 96 items. The /b/-/p/ blocks and /d/-/t/ blocks were alternated. The speakers were randomized across blocks. The items within a block were randomized with the constraint that the items belonging to one pair were never presented consecutively. Different listeners were presented with different randomizations. They were tested in sound-proof booths and were instructed to categorise the first sound of the fragment that they heard as /b/ or /p/ for half of the blocks and as /d/ or /t/ for the other half of the blocks. Before each block started, the two phoneme categories for that particular block appeared on the screen and stayed there during the entire block. Listeners were asked to make their decision by pressing one of two buttons of a response box which corresponded to the phonemes that appeared on the screen. They had to respond within 1.5 seconds. If they failed to do so, the response was not recorded. This occurred in 135 cases in total (0.9%).

Participants

Sixteen volunteers from the Max Planck Institute participant pool were paid to take part in this experiment. All were native speakers of Dutch and none reported any hearing loss. None had taken part in Experiment 2.1.

Results and Discussion

The results of the perception experiment showed that 8.32% of the 15,225 responses did not correspond to the phoneme category written on the list from which the speakers read the items. Inspection of the mismatches showed that there were more mismatching responses to voiced plosives than to voiceless plosives (9.8% for the voiced plosives and 6.8% for the voiceless plosives) and more mismatching responses to labial plosives than to alveolar plosives (8.7% and 7.9% respectively). A binary logistic regression analysis with the number of matching versus mismatching responses as dependent variable and voicing category and place of articulation as independent variables showed that both factors had a significant effect on the number of mismatching responses: Wald(1) = 56.2, $p < .0001$ (voicing), Wald(1) = 10.4, $p < .01$ (place of articulation). The interaction between voicing and place of articulation was also significant (Wald(1) = 33.3, $p < .0001$). The difference between the number of mismatching responses to voiced and voiceless plosives was larger for the labials than for the alveolars (5.9% for /p/ and 11.6% for /b/ versus 7.8% for /t/ and 8.1% for /d/): Wald(1) = 33.3, $p < .0001$.

To find out whether the presence of prevoicing had an influence on the number of mismatching responses, the same logistic regression analysis was conducted on only the voiced plosives. This time, place of articulation and presence of prevoicing were the dependent variables. The effect of prevoicing was significant: Wald(1) = 176.2, $p < .0001$. There were considerably more mismatches to voiced plosives produced without prevoicing than to voiced plosives produced with prevoicing (36.6% versus 1.2%). There was also a significant interaction between prevoicing and place of articulation: Wald(1) = 12.8, $p < .001$. The difference between the percentage of mismatching responses for plosives produced without and with prevoicing was larger for labials than for alveolars: 51.6% (without prevoicing) versus 1.3% (with prevoicing) for the labials, and 25.7% (without prevoicing) versus 1.1% (with prevoicing) for the alveolars.

The identification responses showed that most plosives were perceived as belonging to the voicing category which was intended by the speaker. Although overall the proportion of mismatching responses was very small, analyses showed that there was a difference between the four plosives. The proportion of mismatching responses was largest for the labial voiced plosives, especially for the labial voiced plosives produced without prevoicing. Half of these plosives were perceived as being voiceless. Of the alveolar voiced plosives without prevoicing, a quarter of the tokens were perceived as being voiceless. This suggests that for alveolars the secondary cues are stronger than for labials.

As in Experiment 2.2, two separate CART analyses were conducted for the two places of articulation. This time the purpose was to find out which of the acoustic cues in the signal could best describe the perception of the voicing distinction. In contrast to Experiment 2.2, where the response variable was the voicing category intended by the speaker, the response variable was now the proportion of voiced responses for each token. Since the response variable was now continuous, instead of categorical as in Experiment 2.2, a regression (rather than a classification) tree analysis was performed. The predictor variables of the regression tree analysis were identical to the predictor variables that were used in the classification tree analysis (Experiment 2.2): duration of prevoicing, duration of burst, spectral power of the burst above 500 Hz, SCG of the burst, absolute F_0 immediately after burst offset, F_0 difference, and following phoneme. As before, all tokens with missing measurements were excluded from the analyses (14 bilabials and 18 alveolars).

The two resulting cost-complexity pruned regression trees are shown in Figure 2-10 (labials) and Figure 2-11 (alveolars). The numbers printed directly below each node indicate the mean proportions of voiced responses for all tokens that fell into that final node. Below the mean proportion a small histogram shows the distribution of the proportions of voiced responses for the tokens. As before, the numbers in brackets indicate the numbers of tokens that fell in each final node.

The two regression trees (Figures 2-10 and 2-11) are very similar to the corresponding classification trees of Experiment 2.2 (Figures 2-8 and 2-9). Again, for both the labials and the alveolars, the main split was based on prevoicing duration (9.5 ms and 3.5 ms respectively, as before). The plosives produced with

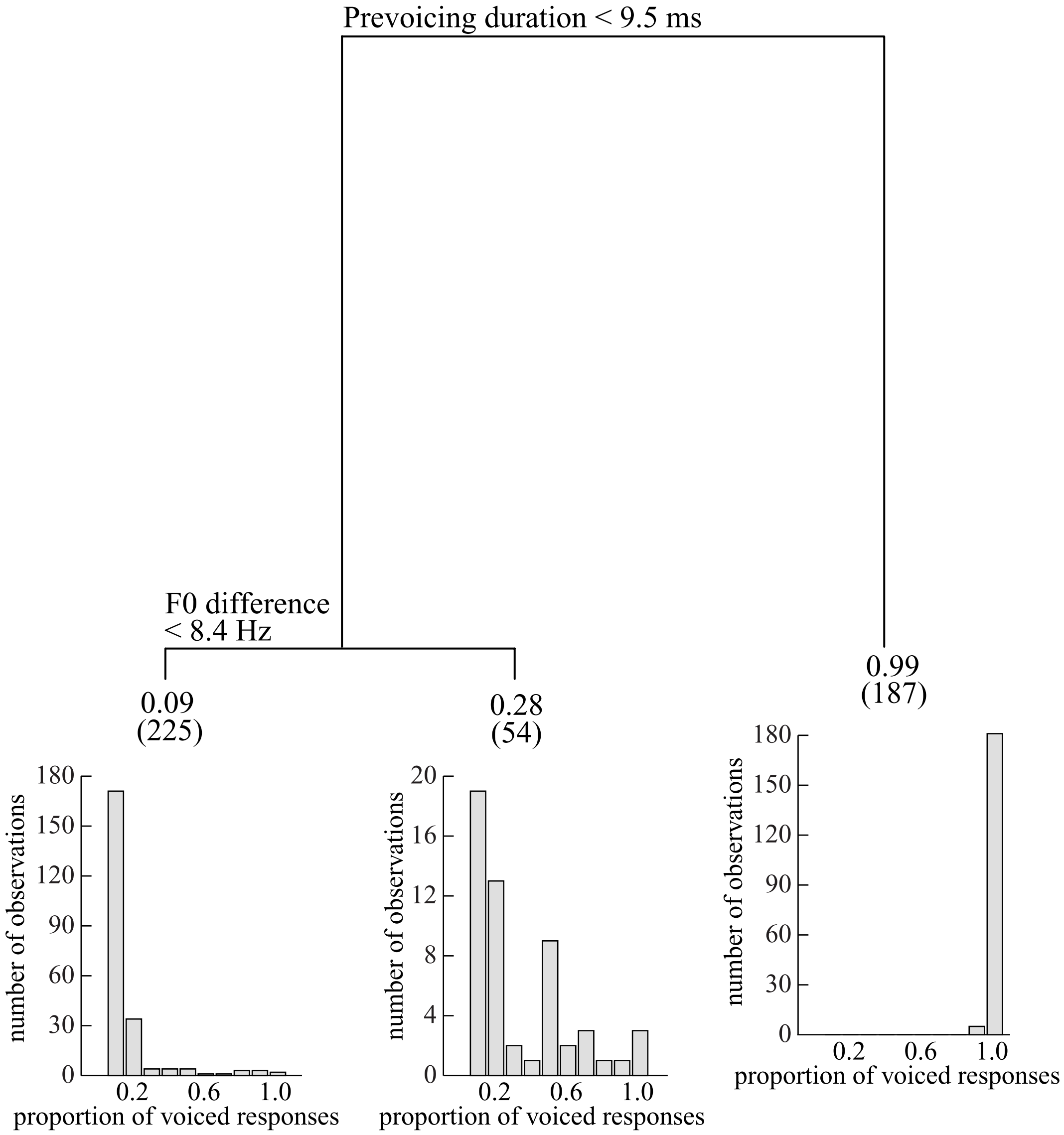

Figure 2-10. *CART analysis of the proportion of voiced responses for the labial plosives. Tokens satisfying the rule printed at the top of each split follow the left branch. Below each final node the mean proportion of voiced responses for the tokens falling in that node is given. The numbers in brackets indicate the numbers of tokens in each node. The histograms show the distributions of the proportions of voiced responses for the tokens that fell in each node.*

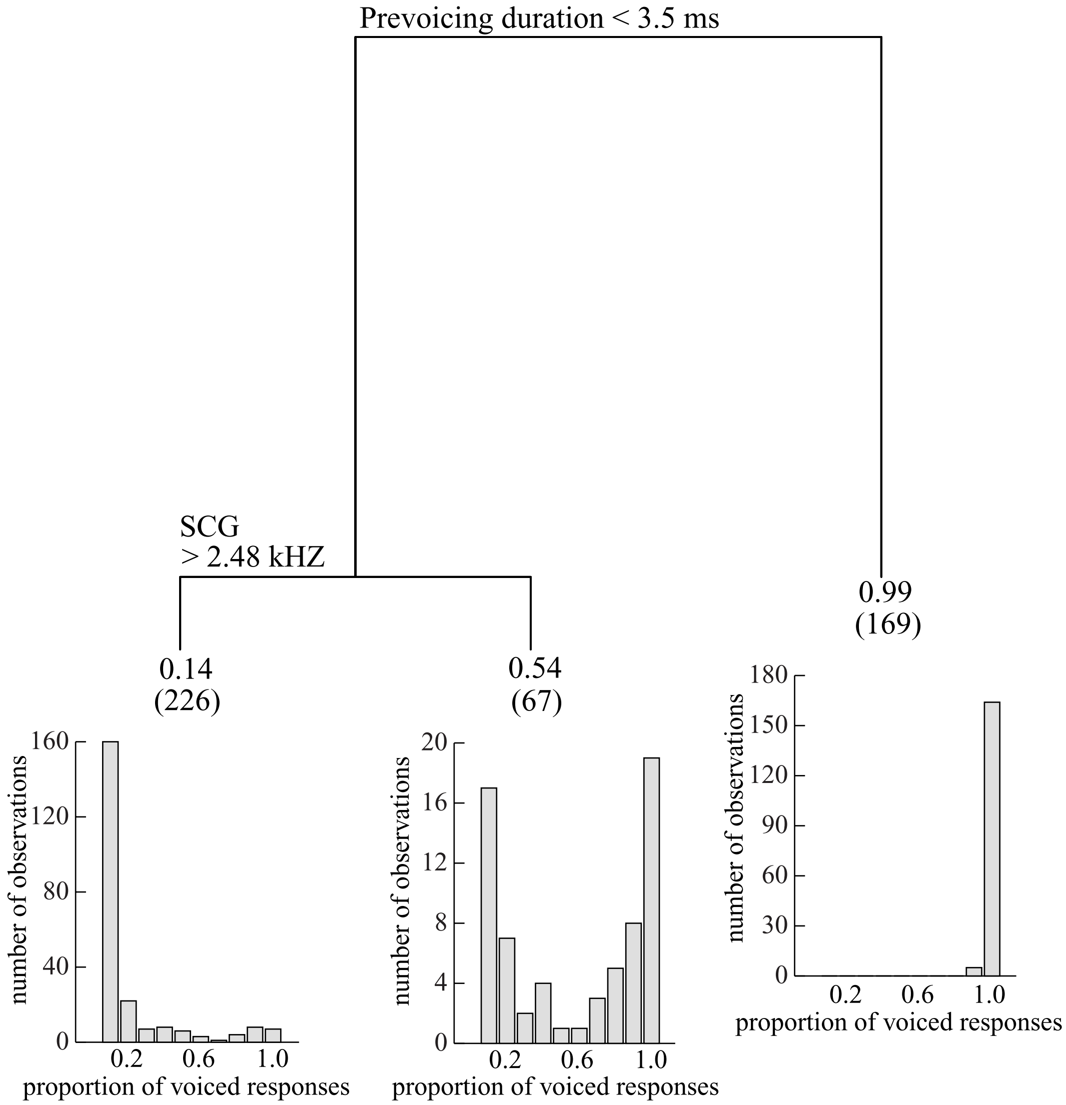

Figure 2-11. *CART analysis of the proportion of voiced responses for the alveolar plosives. Tokens satisfying the rule printed at the top of each split follow the left branch. Below each final node the mean proportion of voiced responses for the tokens falling in that node is given. The numbers in brackets indicate the numbers of tokens in each node. The histograms show the distributions of the proportions of voiced responses for the tokens that fell in each node.*

prevoicing were consistently perceived as being voiced (mean proportion of voiced responses was 0.99 for labials and alveolars), which is also shown in the histograms below the final nodes. The tokens produced without prevoicing were subdivided into two groups. As in Experiment 2.2, the labials were split on the basis of the F_0 movement, while the alveolars were split based on the SCG. Both splits divided the plosives without prevoicing into a relatively large group of plosives for which the mean proportion of voiced responses was small (0.09 for the labials and 0.14 for the alveolars), and into a relatively small group of plosives for which the voicing percept remained more ambiguous (0.28 for the labials and 0.54 for the alveolars). The histograms below these final nodes indicate that for the larger two groups most tokens were indeed perceived as being voiceless while for the two smaller groups there was a lot of variation between the tokens. Most of the alveolar plosives without prevoicing and a low SCG were consistently labeled as voiced or voiceless. There were only a few tokens which were ambiguous. The histogram for the labial plosives without prevoicing and with a F_0 difference larger than 8.4 Hz show a different pattern. Some of the tokens were consistently perceived as being voiceless, while only a few tokens were consistently perceived as being voiced. In addition, there were some tokens which appeared to be fully ambiguous. As we found in the earlier analyses, the strength of the prevoicing cue was much larger than that of the “secondary cues” F_0 difference and SCG. Nevertheless, prevoicing was less dominant for alveolars than for labials. This tallies well with our finding that listeners are more likely to correctly recognize an unprevoiced /d/ than an unprevoiced /b/.

The CART analyses showed that the presence or absence of prevoicing was by far the strongest cue for listeners to identify Dutch initial plosives as voiced or voiceless. This was true for both places of articulation. Both labial and alveolar plosives were perceived as being voiced when produced with prevoicing. The perception of plosives without prevoicing was different for the two places of articulation. Voicing perception in labials produced without prevoicing was influenced by the F_0 movement, while voicing perception of alveolar plosives without prevoicing was influenced by the SCG. These cues could only influence a small subset of the responses and the strengths of these cues were small in comparison to the strength of the main cue prevoicing (indicated by the length of the vertical branches of each split). Interestingly, however, the majority of the voiced plosives produced without prevoicing were still perceived as being voiced on the basis of other cues. These secondary cues were different for the two places of articulation.

SUMMARY AND GENERAL DISCUSSION

This study investigated the production and perception of voicing in Dutch initial plosives. Experiment 2.1 focused on variation in prevoicing, which has been described as the primary cue for initial voiced plosives in Dutch. The productions of 10 different subjects indicated that there was a lot of variation among speakers in terms of number of prevoiced tokens and duration of prevoicing. Five out of 10 subjects prevoiced very consistently, with more than 90% of all their voiced tokens produced with prevoicing. The other five subjects produced prevoicing less frequently, but the proportion of prevoiced tokens varied considerably between those five less frequent prevoicers. Overall, 25% of all tokens produced by all 10 speakers were produced without prevoicing. Several factors appeared to have an effect on prevoicing production. First, the proportion of prevoiced tokens was higher for male speakers than for female speakers. Second, labial plosives were more often produced with prevoicing than alveolar plosives. Third, when the initial plosive was followed by a vowel, prevoicing was produced significantly more frequently and its duration was significantly longer than when the plosive was followed by a consonant.

Experiments 2.2 and 2.3 examined which acoustic properties are produced by speakers to signal the distinction between voiced and voiceless plosives, and which of these are used by listeners when they have to decide whether the plosive is voiced or voiceless. Several durational, spectral and energy cues were measured. All acoustic properties but one had significantly different means for the two voicing categories. A CART analysis showed that of all these acoustic correlates, the presence or absence of prevoicing would be by far the most reliable cue to predict voicing. The tokens that were produced with prevoicing were all assigned to the voiced category. The tokens without prevoicing were further subdivided on the basis of another acoustic property. This property was different for the two places of articulation: the labials without prevoicing were split based on the F_0 difference, while the alveolars without prevoicing were split based on the spectral centre of gravity.

The perception study (Experiment 2.3) showed that the voicing feature of most tokens was perceived as intended by the speaker. Inspection of the mismatching responses showed that most mismatches appeared when the plosive was intended to be voiced but was produced without prevoicing. This suggests that prevoicing plays an important role in perception, which was confirmed by the outcomes of the CART analyses. The analyses showed that prevoicing was by far the strongest cue for the perception of the voicing distinction of both labial and alveolar plosives. Almost all tokens produced with prevoicing were identified as voiced. The majority of the tokens without prevoicing were perceived as voiceless, but a number of unprevoiced tokens were still perceived as voiced. The acoustic cue which most strongly influenced listeners' responses to tokens without prevoicing was different for the two places of articulation. The perception of voicing in labial plosives was influenced most strongly by the F_0 difference, while the perception of voicing in alveolar plosives was influenced most strongly by the spectral centre of gravity. The strength of these cues was fairly low in comparison to the strength of the presence of prevoicing. The correspondence between the analyses of the acoustic and perceptual data was very close.

The results show that the perception of voicing in Dutch plosives is asymmetric: the presence of prevoicing alone provides enough evidence for a listener to be sure that the plosive is voiced, while the absence of prevoicing alone does not provide enough evidence for a listener to categorise the plosive as voiceless. This asymmetry resembles findings in English by Port (1979). He reported that when audible glottal pulsing was maintained through the closure interval, an intervocalic plosive was heard as being voiced, no matter what the values of the other cues (duration of preceding vowel and closure duration) were. It was only when the closure interval was voiceless that the other cues could be effective. Removing all traces of glottal pulsing from the closure interval of a intervocalic /b/, however, did not change the phonemic percept. Only when the duration of the (silent) closure was increased, was the plosive perceived as being voiceless.

It is important to note that the CART analysis was used as a statistical model to analyze the data and not as an explicit model for the listener's behavior. We do not claim that the listener's perceptual system works in the way the CART analysis does, namely by taking into account different cues in serial order. This would require a very complex model involving an extremely rapid succession of low-level decisions, in which listeners would first evaluate prevoicing and would only take other cues into account when this particular cue was absent. Instead, we support models of phonetic categorization in which listeners identify each token by considering in parallel all relevant cues that are available in the speech signal (e.g., Nearey, 1990; Smits, ten Bosch & Collier, 1996). These models claim that listeners first extract a number of perceptually relevant acoustic cues from the speech signal, which together constitute a point in a multidimensional feature space. Associated with this point is a set of probabilities of choosing each of the possible responses, on the basis of which the listener then makes a decision. Thus, the probability that a particular token in Experiment 2.3 belonged to the voiced category would be determined by all relevant cues. The CART analysis showed that the weight of the prevoicing cue was very high. The presence of prevoicing alone brought the probability of a voiced response so close to unity that variation in the other cues had no discernible effect. The absence of prevoicing, on the other hand, did not bring the voiced probability equally close to zero.

When there was no prevoicing, other (weaker) cues, such as the F_0 difference for the labial plosives and the SCG for the alveolar plosives largely determined the class probabilities and therefore the decisions of listeners.

The present study shows that the voicing distinction is acoustically realized differently for labials and alveolars. First, the importance of prevoicing, which clearly plays the most important role in both labial and alveolar plosives, seems to differ between the two places of articulation. Experiment 2.1 and Experiment 2.2 showed that labials were produced more often with prevoicing than alveolars. Experiment 2.3 showed that listeners rely more strongly on prevoicing for labials than for alveolars, because the proportion of mismatching identification responses was larger for labials produced without prevoicing than for alveolars without prevoicing. Alveolar plosives, which have longer and stronger bursts than labial plosives, seem to carry more of the voicing distinction in the burst than labial plosives do. The finding that the difference in the SCG between voiced and voiceless plosives is larger for alveolars than for labials strengthened our impression that the place of articulation for /d/ and /t/ is slightly different. Experiment 2.3 showed that the SCG is an important cue in the perception of alveolar plosives produced without prevoicing but not for labial plosives.

The present study leaves us with an intriguing paradox. Prevoicing is the most reliable cue to the voicing distinction in Dutch initial plosives, yet in a quarter of all voiced plosives prevoicing is absent. Due to the presence of other cues not all voiced plosives without prevoicing were misperceived as voiceless. Nevertheless, although both the production and perception experiments were carried out under optimal conditions, almost 10% of the voiced plosives were mistakenly perceived as voiceless. This proportion is rather high in comparison to identification scores of English voiced plosives (e.g., Smits, 2000), for which the proportion of correct responses was close to 100%. This raises the question whether voicing in plosives is not communicated very accurately in Dutch. The data show that it is not the case that prevoicing is simply difficult to perceive, since all tokens produced with prevoicing were correctly identified as voiced. The voiced tokens which were mistakenly perceived as voiceless, however, were all produced without prevoicing. So the puzzling question is this: given the importance of prevoicing, why do speakers not produce prevoicing more reliably?

The frequent absence of prevoicing may be a peculiarity of the Dutch language. In Dutch, consonants are not always realized according to their underlying [voice]-specifications. Dutch plosives in final position which are underlyingly voiced are always devoiced according to the phonological Final Devoicing Rule (Booij, 1981). Furthermore, many Dutch speakers do not make a distinction between voiced and voiceless fricatives, but produce both categories as voiceless (Cohen, Ebeling, Fokkema & van Holk, 1969; Booij, 1995). The production of voiced plosives without prevoicing in initial position would therefore fit into a pattern in which the underlying [voice]-specification is not always realized phonetically.

A second possible explanation is that the frequent absence of prevoicing is a result of the influence of the English language on Dutch speakers. A similar argument was made by Caramazza and Yeni-Komshian (1974) about the influence of the English language on speakers of Canadian French. They found that in Canadian French prevoicing is frequently omitted, while in European French prevoicing is rarely omitted. Most Dutch speakers speak English as a second language and are daily exposed to the English language. Therefore it is possible that, as in Canadian French, prevoicing is frequently omitted as a result of the presence of a second language in which prevoicing does not play a role in the voicing distinction in initial plosives. Future research on other prevoicing languages and the influence of other language in which prevoicing is not important, should give more insight into this paradox between production and perception.

The present study attempted to give a detailed analysis of the production and perception of Dutch initial plosives in natural speech. Among all different acoustic properties that signal the voicing distinction in these plosives, the presence of prevoicing is by far the strongest cue to the perception of the plosive as belonging to the voiced category. Prevoicing is, however, relatively difficult to realize and is not produced consistently by all speakers. Although the presence of prevoicing signals that the plosive is unmistakably voiced, prevoicing is not a prerequisite for a plosive to be perceived as being voiced. Other acoustic properties can provide sufficient evidence for the plosive to be voiced when prevoicing is absent. These cues are however weak in comparison to prevoicing. Interestingly, the voicing distinction in alveolar plosives seems to be realized with a small difference in the place of articulation, which makes the voicing distinction in alveolar plosives more robust than in labial plosives when prevoicing is absent.

PREVOICING IN DUTCH INITIAL PLOSIVES: PRODUCTION, PERCEPTION, AND WORD RECOGNITION

CHAPTER 3

Petra M. van Alphen (under revision), to appear in E. J. van der Torre & J. van de Weijer (Eds.), *Voicing in Dutch*

INTRODUCTION

This chapter focusses on the phonological voicing distinction in Dutch initial plosives, that is, the distinction between [+voice] and [-voice] in those plosives. Although most languages contrast these two phonemic classes (which I will refer to as *voiced* and *voiceless* plosives), the way in which this phonological distinction is implemented phonetically varies across languages. Lisker and Abramson (1964) investigated eleven languages and measured the time between the onset of vocal fold vibration and the release of a plosive, which they referred to as Voice Onset Time (VOT). They established that, across languages, VOT is essentially tri-modal. The three categories based on VOT were: plosives with a negative VOT, produced with a voiced lead (i.e., with voicing during the closure); plosives with a slightly positive VOT, produced with almost no aspiration; and plosives with a clear positive VOT, produced with aspiration.

Some languages, such as Thai, employ all three modes in a three-way voicing distinction. Most languages, however, have a two-way voicing distinction, which is implemented by two adjacent modes, one of which is associated with the voiced, and the other with the voiceless plosive. Keating, Linker and Huffman (1983) surveyed 51 languages and observed that almost all these languages use at least some kind of voiceless unaspirated plosive and that of the two categories contrasting the voiceless unaspirated plosive, fully voiced and voiceless aspirated plosives are about equally common. Furthermore, they observed that within many languages there is variation across positions in a word. In the present chapter I will focus on plosives in initial position.

Germanic languages such as Danish, English and German contrast voiceless unaspirated and voiceless aspirated plosives in initial position (Keating, 1984). Dutch, however, is unusual among Germanic languages in that it does not include this contrast. Instead, Dutch, along with other languages such as Arabic, Bulgarian, French, Japanese, Polish, Russian and Spanish, has a traditional voicing contrast (Keating, 1984; Lisker & Abramson, 1964). That is, the voiced plosives are produced with a voice lead, which I will refer to as *prevoicing*, and the voiceless plosives are produced with little or no aspiration. Figure 4-1 shows an example of a voiced plosive with prevoicing and an example of a voiceless plosive. There are three plosives in Dutch which belong to the voiceless category, namely [p], [t], and [k], while there are only two plosives which belong to the voiced category, namely [b] and [d]. The velar voiced plosive [g] only occurs in loan words and is therefore not discussed here.

In this chapter I will first describe the production of prevoicing and its occurrence in Dutch and I will then focus on the role of prevoicing in perception. There seems to be an interesting paradox between production and perception: prevoicing is frequently absent in Dutch initial voiced plosives, but the presence of prevoicing is nevertheless a very strong cue for the perception of voicing in these plosives. In order to fully understand the influence that prevoicing has on perception it is important not only to study phoneme perception, but also to study word recognition. Words are after all the meaningful units which a listener has to recognize in order to retrieve the message of the speaker. Two priming experiments will be presented which investigate the effects of two types of prevoicing variation on word recognition. These experiments, and others that I will review, lead to the conclusion that word recognition is sensitive to prevoicing variation, but only to the type of variation that is relevant for the distinction between lexical candidates.

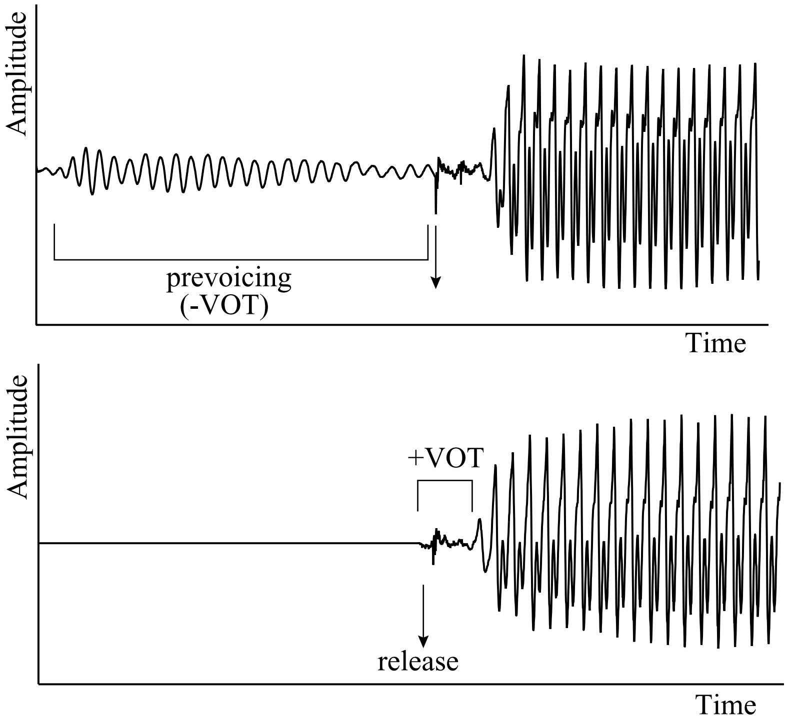

Figure 3-1. *Waveforms of the initial voiced plosive and part of the vowel of the Dutch word /bo:t/ (upper panel) and of the initial voiceless plosive and part of the vowel of the Dutch word /po:t/ (lower panel)*

PRODUCTION OF PREVOICING

Prevoicing refers to the presence of vocal fold vibration during the closure of the plosive. According to the myoelastic-aerodynamic theory of phonation (van den Berg, 1958), the vocal folds will vibrate only when they are properly adducted and tensed, and when there exists a sufficient transglottal pressure gradient to result in a positive airflow through the glottis from the lungs. When a vowel or continuant consonant is produced it is not very difficult to obtain sufficient transglottal pressure: the vocal tract is open and therefore the subglottal pressure will be higher than the supraglottal pressure as long as there is sufficient air in the lungs. During the production of a plosive, however, all out-going airways are closed. The blocking of out-flowing air, causes the supraglottal pressure to increase rapidly, which results in a rapid decrease of the transglottal pressure. It is therefore relatively difficult to let the vocal folds vibrate during the closure. As the volume above the glottis increases, a sufficient transglottal pressure gradient can be obtained for a longer period of time, since the supraglottal pressure increases less rapidly. Enlargement of the supraglottal cavity will thus help to initiate and maintain voicing. This enlargement can be obtained actively, by lowering the larynx, raising the soft palate, advancing the tongue root, or drawing the tongue dorsum and blade down (see Westbury, 1982), or passively when the walls of the supraglottal cavity are lax which allows them to expand in response to the internal pressure (Rothenberg, 1968).

Children acquire the production of prevoicing relatively late (Kewley-Port & Preston, 1974), which also suggests that prevoicing production is relatively difficult. Nevertheless, studies on the production of prevoicing in languages such as Polish (Keating, Mikos & Ganong, 1981), Lebanese Arabic (Yeni-Komshian, Caramazza & Preston, 1977) and European French (Caramazza & Yeni-Komshian, 1974) show that adult speakers rarely omit prevoicing when producing voiced plosives. Only one study, on Canadian French (Caramazza & Yeni-Komshian, 1974) has found a substantial degree of overlap between the VOT distributions of voiced and voiceless plosives; no less than 58% of the voiced tokens in that sample (N=90) were produced without prevoicing. Caramazza and Yeni-Komshian argued that in Canadian French the VOT values are shifting as a result of the influence of Canadian English.

Until recently there were no studies which systematically investigated the occurrence of prevoicing in Dutch. It is, however, important to know how prevoicing varies naturally in order to understand effects of prevoicing variation on speech perception. The way in which the speech recognition system treats a particular acoustic property largely depends on the variation in the occurrence of this property in natural speech. In one of my recent studies (van Alphen & Smits, submitted) the occurrence of prevoicing in Dutch was therefore investigated.

Van Alphen and Smits (submitted; see Chapter 2) asked ten Dutch speakers to produce 32 real words and 32 nonsense words with initial voiced plosives (/b/ or /d/). These items were presented randomly in a list with fillers (including the same items starting with voiceless plosives), such that the listener's attention was not drawn to the voicing distinction. The results showed that 25% of the tokens with initial voiced plosives were produced without prevoicing. The proportion of prevoiced tokens was found to be influenced by the following factors: sex of the speaker (male or female), place of articulation of the plosive (labial or alveolar), and the phoneme following the plosive (vowel or consonant). All these factors might have an effect on the vocal tract volume or on the extent to which the vocal tract can be expanded. The smaller the volume of the vocal tract, the faster the supraglottal pressure increases and the more difficult it is to produce prevoicing. Male speakers are expected to have a larger vocal tract size than female speakers, which makes it easier for males to produce prevoicing. In line with this expectation, male speakers produced prevoicing more often than female speakers (86% versus 65%). The place of articulation of a plosive was expected to influence the extent to which the vocal tract can be expanded passively due to raised intraoral pressure. For dental stops, the pharyngeal walls and part of the soft palate can yield to expansion of the oral cavity, while for labial stops these surfaces plus all of the tongue surface and parts of the cheek can participate in the expansion (Houde, 1968; Rothenberg, 1968). The oral cavity can thus be expanded more during the production of labial plosives than during the production of dental plosives. Van Alphen and Smits indeed found that labial plosives were more often produced with prevoicing than alveolars (79% versus 72%). Finally, the following phoneme was expected to affect the vocal tract size and the extent to which the different mechanisms (passively and actively) could expand the vocal tract size, and thus the proportion of prevoiced tokens. No effect of vowel height was found, but plosives followed by a vowel were more often prevoiced than plosives followed by a consonant (86% versus 65%).

Although it seems that prevoicing was absent in the cases where the aerodynamics made it harder to produce prevoicing, it can not be the case that prevoicing is simply too difficult to produce in particular cases, since other studies on prevoicing production in other languages (e.g., Keating, Mikos & Ganong, 1981; Yeni-Komshian, Caramazza & Preston, 1977; and Caramazza & Yeni-Komshian, 1974) did not find such a large proportion of unprevoiced tokens. This suggests that Dutch speakers make less effort to produce prevoicing, resulting in a relatively large proportion of voiced plosives without prevoicing, especially in the cases in which it is difficult to produce prevoicing. We can only speculate about the reason for this. It may be the case that the way in which the voicing distinction in Dutch is implemented phonetically is changing as a result of the influence of English on the Dutch language.

THE ROLE OF PREVOICING IN THE PERCEPTION OF THE VOICING DISTINCTION

Now that we know that prevoicing is frequently absent in Dutch initial voiced plosives, we can ask what influence this has on perception. Are the voiced tokens produced without prevoicing still perceived as voiced? In other words, is the production of prevoicing essential for the plosives to be perceived as voiced, or are other acoustic cues present and strong enough to evoke a voiced percept? We know from the previous literature that VOT is not the only acoustic property which covaries with the voicing distinction in plosives (see for example Jessen, 1998 for German; Slis & Cohen, 1969 for Dutch). Van Alphen and Smits therefore examined what other acoustic properties were present in the acoustic realizations of Dutch initial plosives which could serve as potential perceptual cues to the voicing distinction. The following six measures were obtained from a sample of 480 voiced tokens and 480 voiceless tokens: duration of prevoicing, duration of the burst, power of the burst, spectral centre of gravity of the burst, F_0 immediately after burst offset, and F_0 movement into the vowel (see van Alphen & Smits for a detailed description of the measurements). Except for the F_0 immediately after burst offset, all measures showed a significant difference between voiced and voiceless plosives. In addition to the finding that voiced plosives had more prevoicing than voiceless plosives (which were never produced with prevoicing) the data showed that all three measures involving the burst (the duration, power and spectral centre of gravity) were lower for voiced than for voiceless plosives. Finally, the mean F_0 difference (that is, the difference between the F_0 in the middle of the following vowel and the F_0 immediately after the burst) was positive for tokens starting with voiced plosives, consistent with a rising F_0, while it was negative for tokens starting with a voiceless plosive, consistent with a falling F_0. These differences indicate that the speech signal contains a variety of potential perceptual cues for the voicing distinction.

Sixteen listeners were then asked to identify the 960 tokens as voiced or voiceless. Regression tree analysis of the responses indicated that, of all measured acoustic properties, the presence or absence of prevoicing was by far the strongest cue to the voicing distinction as perceived by listeners. All tokens produced with prevoicing were perceived as voiced. Tokens without prevoicing, however, were perceived either as voiced or voiceless. In those cases, the perceived voicing category depended on the value of the other acoustic cues in the signal. When those cues were in favor of the voiced category, the tokens were perceived as voiced, despite the absence of prevoicing. The acoustic cue which most strongly influenced listeners' responses to tokens without prevoicing was different for the two places of articulation. The perception of voicing in labial plosives was influenced most strongly by the F_0 difference from the burst of the plosive into the vowel: a higher F_0 difference yielded a higher proportion of voiced responses. The perception of voicing in alveolar plosive appeared to be influenced most strongly by the spectral centre of gravity of the spectral noise of the burst: a higher spectral centre of gravity yielded a lower proportion of voiced responses. Nevertheless, these secondary cues were rather weak in comparison to prevoicing. Of all tokens produced without prevoicing which were intended to be voiced, 37% were identified as voiceless. The absence of prevoicing clearly decreases the probability that a token is perceived as voiced.

The results of the study by van Alphen and Smits indicate that the presence or absence of prevoicing plays an important role in the phonetic realization and the perception of the phonological voicing distinction in Dutch initial plosives. The role of prevoicing is, however asymmetric: voiceless plosives are always produced without prevoicing, while voiced plosives are not always produced with prevoicing. In line with this, tokens produced with prevoicing are always perceived as voiced, while tokens produced without prevoicing are not always perceived as voiceless.

So far, I have argued that prevoicing has a strong influence on the identification of Dutch initial plosives as voiced or voiceless. Of course, speech perception involves more than the perception of phonological features or the perception of single phonemes. The core process in speech perception is the recognition of words. Since words are the units which convey meaning, the recognition of words is an essential component of how the listener retrieves the message of the speaker. Thus, the next step one has to take in order to fully understand the effect of prevoicing variation on speech perception is to examine the influence of prevoicing variation on the recognition of words.

EFFECTS OF FINE-GRAINED ACOUSTIC DETAILS ON WORD RECOGNITION

Word recognition involves the mapping of the speech signal onto stored lexical knowledge. As the utterance unfolds over time, multiple lexical candidates are activated as a result of the acoustic input. The activation of a lexical candidate at a particular moment in time reflects the goodness of fit with the available acoustic input at that moment. The candidate that eventually matches the acoustic input best will be recognized. It appears that the activated lexical candidates compete with each other for recognition; the most strongly activated candidate will suppress the activation of the other lexical candidates and win the competition (see McQueen, in press, for an overview of the evidence for the existence of competition between lexical candidates).

The speech signal is highly variable, however, and not all acoustic information is relevant for the recognition of words. Therefore, the assumption is that listeners perform a detailed phonetic analysis of the acoustic input prior to lexical access. At the prelexical level, the incoming speech signal is normalized and useful information is extracted from the speech signal and translated into abstract representations. Many different units have been proposed as prelexical representations, including syllables (Mehler, 1981), semi-syllables (Massaro, 1987), phonemes (Foss & Blank, 1980; Nearey, 2001), allophones (Luce, Goldinger, Auer & Vitevitch, 2000) and features (Stevens, 2002). So far, research has not provided us with conclusive evidence singling out one of these units.

These prelexical representations, whatever their exact nature, are assumed to activate word representations. The prelexical level acts thus as an intermediate level at which the speech signal is analyzed and filtered. On this account, it is important to distinguish between acoustic detail that is normalized away at the prelexical level and that which is passed on to the lexical level. It could be the case that at the prelexical level discrete decisions are made (for example, hard phoneme decisions: Is this a [b] or is this a [p]?), and that most acoustic detail is thus normalized away. In contrast, it could also be the case that the prelexical level preserves part of the acoustic detail such that the output of the prelexical level is graded (for example, one particular token of the labial plosive results in more activation of the prelexical representation for [b] than another token does). In other words, how much acoustic detail is still present in the information that reaches the lexical level?

Many studies have shown that lexical activation is in fact sensitive to fine-grained acoustic information (see McQueen, Dahan & Cutler, in press, for a detailed overview). For example, Spinelli, McQueen and Cutler (2003) showed that French listeners are sensitive to small durational differences in the consonants of two utterances, such as in the final [ʁ] of *dernier* in *dernier oignon* (last onion) and the final [ʁ] of *rognon* in *dernier rognon* (last kidney). Note that these two phones are phonemically identical due to liaison (i.e., the appearance of the final [ʁ] of *dernier*) in *dernier oignon*. Spinelli at al. constructed a cross-modal identity priming experiment in which listeners had to perform lexical decisions on visual targets such as *oignon* or *rognon*, which were preceded by spoken versions of either *dernier oignon* or *dernier rognon*. A significant priming effect was found (that is, faster lexical decisions to targets preceded by identical primes in comparison to lexical decisions to targets preceded by unrelated primes), but only when the target matched the speaker’s intention (e.g., when the target *rognon* was preceded by the utterance *dernier rognon* rather than when it was preceded by the utterance *dernier oignon*). The consonants in the liaison environment appeared to be shorter than the word-initial consonants (e.g., the [ʁ] in *dernier oignon* was about 10 ms shorter than the [ʁ] in *dernier rognon*). The results of this identity priming study thus showed that the speech recognition system in French listeners is sensitive to very small durational differences.

Studies such as Spinelli et al. (2003) show that the degree of activation of lexical candidates is influenced by fine-grained differences in the speech signal and thus suggest that small acoustic details are preserved by the prelexical level and can reach the lexicon. In other words, they challenge the view that discrete decisions (for example phonemic decisions) are made at the prelexical level. It seems that graded activation of prelexical representations is passed on continuously to the lexical level. This is in line with spoken-word recognition models such as TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994; Norris, McQueen & Cutler, 2000), in which information flows continuously from a prelexical level of processing to the lexical level.

Among the studies which report effects of fine-grained acoustic details on lexical activation, there are a number of studies which focus on variation in VOT. Andruski, Blumstein and Burton (1994) obtained variations in English VOT by removing one third or two thirds of the original positive VOT of voiceless plosives which appeared word initially. They examined the influence of these VOT variations on the activation of lexical candidates in a within modality associative priming experiment. In this experiment, listeners were asked to perform a lexical decision task on spoken targets which were preceded by spoken primes. A target word, for example *queen*, was preceded by either a semantically unrelated prime, such as *bell*, or by a semantically related prime, such as *king*. All related primes started with a voiceless plosive and appeared in three different VOT conditions: with unaltered VOT, with two thirds of the original VOT, or with one third of the original VOT. Furthermore, half of the related primes were words which had a voiced word competitor, that is, changing the initial voiceless plosive into the matching voiced plosive resulted in a word, for example *pear* (*bear)*. The other half of the primes were words which did not have a lexical competitor, for example *king (ging)*. The reaction time (RT) patterns of the lexical decisions showed that listeners were faster to make lexical decisions to targets when they were preceded by related primes than when they were preceded by unrelated primes. Interestingly, lexical decisions to targets preceded by the primes with one third of the original VOT were significantly slower than lexical decisions to the same targets preceded by primes with unaltered VOT. The presence of a voiced word competitor seemed not to influence these effects. Furthermore, these effects of VOT manipulation only appeared when the delay between the offset of the target and onset of the prime was short (50 ms); they did not appear when the delay was longer (250 ms).

Utman, Blumstein and Burton (2000) explored the influence of similar VOT differences on lexical activation using a uni-modal identity priming experiment. This time, both words and nonwords starting with voiceless plosives were used as primes. Spoken targets were preceded by the same natural tokens of those targets, or by tokens in which the VOT was shortened. The findings for the word primes were consistent with the findings by Andruski et al. (1994): lexical decisions to spoken word targets, such as *kiss,* were slower when these targets were preceded by spoken primes, such as *kiss,* of which only one third of the original VOT was preserved, than when these targets were preceded by primes which were identical (with unaltered VOT). When targets and primes were nonwords, however, no effect of the VOT reduction was found on the lexical decisions.

McMurray, Tanenhaus and Aslin (2002) also investigated the effect of VOT variation on lexical access in English. In an eye-tracking experiment, listeners were seated in front of a screen on which four pictures appeared, of which two were members of a minimal word pair like *beach-peach*. Meanwhile they heard repetitions of each of the nine steps of a word-word continuum, which varied along the VOT dimension (e.g., from *beach* to *peach*). McMurray et al. found that the mean proportion of fixations to the two target pictures varied gradually as a function of the VOT of the initial plosives.

These experiments show that differences in English positive VOT are not normalized away at the prelexical level, but that this type of acoustic detail is passed on to the lexical level where it can affect the degree of lexical activation. Can similar effects be observed for differences in the negative VOT of initial plosives in Dutch? This question was addressed in the priming experiments presented below. In order to understand the predictions which were made for Dutch, however, it is important to first consider the differences between VOT in English and Dutch.

Although in both English and Dutch VOT plays an important role in the phonological voicing distinction of word-initial plosives, the phonetic realization of voiced and voiceless plosives is rather different in the two languages. While in English the informative value of VOT lies in the positive VOT range, that is, in the exact duration of aspiration, in Dutch it is the presence or absence of prevoicing which seems to be important (van Alphen & Smits, submitted). In English, the phoneme boundary between voiced and voiceless plosives in terms of VOT is not fixed, but varies on a continuous scale, for example as a function of speaking rate (Green & Miller, 1985; Summerfield, 1981). English listeners are therefore required to make fine temporal distinctions along the VOT dimension in order to perceive the plosive as voiced or voiceless. In contrast, Dutch listeners do not need to establish the exact duration of the VOT to perceive the voicing distinction, since, as described above, the voicing distinction in Dutch is signalled by the presence or absence of prevoicing, rather than by the exact amount of prevoicing. Similar suggestions have been made by Keating, Mikos and Ganong (1981) about the comparison between the informational value of VOT variation for English versus Polish listeners.

Given these differences between VOT in English and Dutch, the first prediction is that as long as Dutch initial plosives have prevoicing, differences in the exact amount of prevoicing will not affect lexical activation. After all, the exact duration of prevoicing will not help listeners to distinguish between two alternative lexical candidates such as *beer* (bear) or *peer* (pear). Therefore, this type of uninformative acoustic detail should be normalized away at the prelexical level. As a result, shortening prevoicing duration should not affect lexical activation. The difference between the presence or absence of prevoicing, however, does carry information about the Dutch voicing distinction. Recall that van Alphen and Smits (submitted) showed that the absence of prevoicing decreased the probability that that token was voiced. Therefore, the second prediction is that the deletion of prevoicing would affect lexical access.

EXPERIMENT 3.1

To test these predictions, three prevoicing values were chosen (0, 6 and 12 periods of prevoicing) such that the smallest duration was zero and such that the physical difference between the subsequent prevoicing durations was the same. Importantly, all three degrees of prevoicing fell within the natural range of prevoicing duration as established by van Alphen and Smits (submitted). The expectation was to find an effect of the difference between the absence and presence of prevoicing (0 versus 6 periods of prevoicing), but not of prevoicing shortening (12 versus 6 periods of prevoicing). Furthermore, the experiments explored whether a possible effect of prevoicing differences would be influenced by the frequency of the prime words.

Following Andruski et al. (1994), the associative priming task was chosen. But primes and targets were presented in different modalities (spoken primes were followed by visual targets), rather than within one modality. The reasons for this were twofold. First, Andruski et al. only observed effects when the delay between the offset of the prime and onset of the target was short (50 ms). Since the VOT manipulation in Dutch appeared even earlier in the prime word than in English (prevoicing appears at the beginning of the plosive while aspiration appears after the burst of the plosive), it seemed preferable to present the targets immediately after the offset of the spoken prime. If both prime and target were presented auditorily with zero delay, the prime could mask the end of the target. The cross-modal version of the associative priming task avoids this problem. Second, the use of a visual target ensured that what was tested was activation at the lexical level, rather than activation at the prelexical level as a result of possible phonological overlap between prime and target. Although most primes and targets did not overlap phonologically (for example, *bloem* (bloem) - *roos* (rose)) some of the primes and targets did (for example, *brood* (bread) - *boter* (butter)).

The underlying idea in the use of the associative priming task is that the processing of a stimulus (the prime) may facilitate the subsequent processing of a following stimulus (the target) if the prime is semantically related to the prime. To measure the influence of the presentation of the prime on the processing of the target, participants are asked to perform a task such as lexical decision on the targets. The RTs of these decisions are then compared to the RTs in a baseline condition, in which the target is preceded by a semantically unrelated prime (see, for example, Marslen-Wilson & Zwitserlood, 1989).

If it is indeed the case that deletion of prevoicing affects lexical activation while differences in the amount of prevoicing do not, the following patterns should be observed: faster lexical decisions should be made to targets such as *roos* (rose) when the preceding semantically related prime *bloem* (flower) starts with prevoicing than when the same prime has no prevoicing; and there should be no difference between lexical decisions to targets preceded by related primes with 12 periods of prevoicing and those to targets preceded by primes with 6 periods of prevoicing. If it is the case, however, that the prelexical level does not normalize away the difference in prevoicing duration (12 versus 6 periods of prevoicing) such that this type of variation does affect lexical activation, different priming effects should be found for primes with 12 and 6 periods of prevoicing. The expectation would then be that primes such as *bloem* starting with 12 periods of prevoicing will result in stronger activation of the lexical representations of those words (e.g., the lexical representation of *bloem)* than the same primes starting with 6 periods with prevoicing, since plosives with 12 periods of prevoicing are further away from the phoneme boundary.

Method

Participants

Forty-eight students were paid to take part in the experiment. None of them reported any hearing loss.

Materials

Two types of words were selected as primes: 40 high frequency words (HF words), and 40 low frequency words (LF words). The mean frequency of the HF words was 97 per million words and the mean frequency of the LF words was 2 per million words (from the CELEX lexical database, Baayen, Piepenbrock & Gulikers, 1995). Half of the HF words started with a /b/ and the other half started with a /d/. Of the LF words, 25 words started with a /b/ and 15 with a /d/. All words were mono- or bisyllabic; the bisyllabic words all had a strong-weak stress pattern.

For each word a semantically related word was selected to serve as a visual target. This was done by asking 23 subjects to give their associations for each word. An associated word was regarded as a good target when the word was given in response by more than 25% of the subjects and when the difference between that associated word and the next most frequent associated word was greater than 10%. The mean frequency of the targets associated with the HF prime words was 133 per million words and the mean frequency of the targets associated with the LF prime words was 42 per million words. For each target an unrelated prime was also chosen which matched the related prime in length and started with the same phoneme (/b/ or /d/). In addition to the 80 word targets there were 40 nonword targets preceded by (unrelated) primes starting with a voiced plosive (half of them started with a /b/ and half of them with a /d/). Furthermore, 200 other targets were paired with primes that started with a phoneme other than a /b/ or /d/: 120 nonword targets with unrelated primes; and 80 word targets, of which 20 were preceded by a related prime and 60 by an unrelated prime. The design is summarized in Table 3-1 and all materials are given in Appendix 3-A.

Stimulus construction

All primes were recorded several times on digital audio tape (at a sampling rate of 48 kHz with 16-bit resolution) by a male native speaker of Dutch. The utterances were then digitized at a sample rate of 16 kHz. For the three prevoicing priming conditions and the unrelated condition, tokens were chosen which were produced clearly and with prevoicing. Subsequently, the original prevoicing of each related prevoiced item was replaced by 12, 6 or 0 periods of prevoicing (corresponding to 129, 64 or 0 ms of prevoicing for /b/ and to 127, 62 and 0 ms of prevoicing for /d/), in order to create the three different prevoicing conditions. The first full period of prevoicing plus the lead-in (of 5 ms) of a natural token of the word /bʉs/ (bus) was chosen as the first period of prevoicing for the two conditions with prevoicing for the items starting with a labial plosive. Similarly, the last prevoicing period of that same token of /bʉs/ always served as the last prevoicing period in these two conditions. The intervening prevoicing periods (10 or 4) were randomly chosen from the /bʉs/ token. The same procedure was applied to create the prevoicing 12 and prevoicing 6 conditions for the items starting with

Table 3-1. *Design of Experiments 3.1 (associative priming) and 3.2 (identity priming). For each combination of priming condition and prime frequency (including nonword primes) examples of a prime and target are given. Real words have their English translation in parentheses.*

		Prime Type		
	Priming Condition	HF words	LF words	Nonwords (Exp 3.2 only)
PRIME	Prevoicing 12	bloem (*flower*)	beits (*stain*)	breld
	Prevoicing 6	bloem (*flower*)	beits (*stain*)	breld
	Prevoicing 0	bloem (*flower*)	beits (*stain*)	breld
	Unrelated	baan (*job*)	broche (*brooch*)	biem
TARGET	Experiment 3.1 (associative)	roos (*rose*)	verf (*paint*)	-
	Experiment 3.2 (identity)	bloem (*flower*)	beits (*stain*)	breld

an alveolar plosive, but now the prevoicing periods were derived from a natural token of the word /dʉs/ (thus). To control for any splicing effects, the prevoicing of each of the unrelated primes was also replaced by six periods of prevoicing.

Procedure

Primes were presented binaurally over headphones in a sound-damped booth. Immediately after the offset of the prime the visual target was presented in lower case on a computer screen. Subjects were instructed to listen to the word and then decide as quickly as possible whether the stimulus on the screen was a word or a nonword, by pressing one of two buttons. Four lists were constructed with priming condition counterbalanced across lists. Each subject therefore saw each target only once, preceded by one of the four possible primes for that item. Furthermore, the lists contained all fillers such that half of the targets were words and the other half nonwords. Of the total of 320 pairs in a given list, 80 pairs (25%) were related.

After the associative priming experiment, all test items that were used as related primes were presented to the same listeners for identification of the initial phoneme. In addition to the 240 word tokens starting with a voiced plosive (/b/ or /d/), the identification task contained three repetitions of 80 distractor words starting with a voiceless plosive (/p/ or /t/). The items were blocked by place of articulation. Half of the subjects started with the labial plosives and half of the subject started with the alveolar plosives.

Results and Discussion

The results of the phoneme identification task showed that, overall, 97% of the items starting with a voiced plosive were identified as voiced. One item appeared to be misrecorded and was therefore removed from all further analyses. Table 3-2 shows the percentage of voiced responses in each of the three prevoicing conditions for HF words and LF words separately.

The proportions of voiced responses were converted through an arcsine transformation (Studebaker, 1985) and submitted to repeated measures subjects (F1) and items (F2) analyses of variance (ANOVAs) with the factors frequency and prevoicing. There was a main effect of prevoicing ($F1(2,94) = 47.89$, $p < .001$; $F2(2,154) = 38.46$, $p < .001$). No other effects were significant. Tukey honestly significant difference (HSD) tests showed that the proportion of voiced responses to tokens without prevoicing was significantly lower than those to items with prevoicing (either 12 or 6 periods of prevoicing). Nevertheless, these tokens without prevoicing were in general still perceived as voiced. Note that all items starting with voiced plosives were words, which could have biased listeners to respond with the voiced category. Inspection of the RTs of the identification responses suggested that some of the responses were initiated even before the end of the prevoicing (in particular when the plosive started with 12 periods of prevoicing). Apparently, in some cases the presence of prevoicing alone provided sufficient information that the plosive was voiced. Since not all responses were initiated after the end of the prevoicing, it was not possible to correct for the length of the prevoicing. Therefore, there was no accurate way to analyze the RTs of the identification data in this study or in any of the following experiments.

Table 3-2. *Percentage of voiced responses in the identification task of Experiments 3.1 and 3.2*

		HF words	LF words	Nonwords
Experiment 3.1	Prevoicing 12	98.0	98.3	-
	Prevoicing 6	98.6	98.5	-
	Prevoicing 0	93.2	93.4	-
Experiment 3.2	Prevoicing 12	98.7	98.6	99.3
	Prevoicing 6	98.9	99.4	99.4
	Prevoicing 0	93.8	93.5	82.9

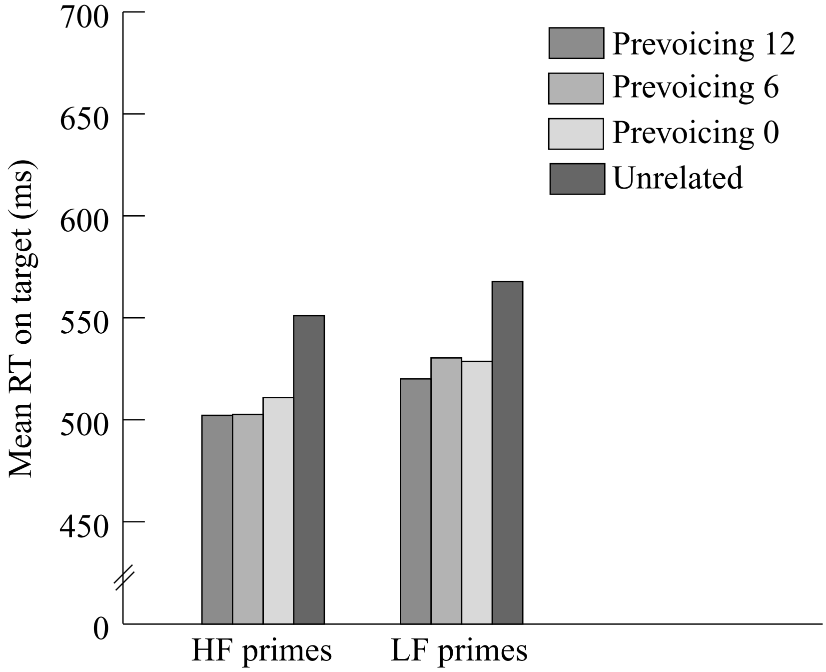

Figure 3-2. *Mean reaction times (RTs) to word targets preceded by high frequency (HF) and low frequency (LF) primes in each of the four priming conditions in Experiment 3.1 (associative priming)*

In the associative priming study the effect of the different prevoicing durations was investigated by measuring lexical decision RTs to the visual targets. RTs were measured from target onset and therefore there was no need to correct for differences in the duration of the prime as a result of the prevoicing manipulation. The mean latencies of correct lexical decisions to word targets are shown in Figure 4-2. Subjects showed semantic facilitation, responding faster to targets preceded by semantically related primes than to targets preceded by unrelated primes. Repeated-measures subjects (F1) and items (F2) ANOVAs with prime type (12 periods, 6 periods, no prevoicing, unrelated), frequency (HF and LF), phoneme (/b/ and /d/) as factors showed significant effects of prime type: F1(3,141) = 29.41, $p < .001$; F2(3,225) = 24.60, $p < .001$, and of frequency: F1(1,47) = 33.13, $p < .001$; F2(1,75) = 6.33, $p < .05$. No other effects were significant.

In addition, t-tests on the following three planned comparisons were carried out: prevoicing 12-prevoicing 6, prevoicing 6-unrelated, and prevoicing 0-prevoicing 6. The outcomes of the two-tailed t-tests showed that the difference between the prevoicing 6 condition and the unrelated condition was significant (t1(47) = -6.39, $p < .001$; t2(78) = -5.82, $p < .001$), but that the other two differences were not significant. This indicates that lexical decisions were significantly faster when the target was preceded by a semantically related prime than when the target was preceded by a semantically unrelated prime, and that the lexical decisions latencies were not affected by the degree of prevoicing. There were also no differences among the error rates of the three prevoicing conditions.

The frequency effect indicated that RTs to targets preceded by a HF prime were faster than RTs to targets preceded by a LF prime (517 ms versus 537 ms). RTs were negatively correlated with target word frequency ($r(79) = -0.249$, $p < 0.05$, two-tailed), but were not correlated with prime frequency, showing that the frequency effect on RTs was caused by target frequency, not prime frequency.

These results suggest that differences in prevoicing duration of plosives do not influence lexical access. It is possible, however, that the VOT variation tested here does influence lexical access but that the associative priming task is not sensitive enough to measure an influence of such small acoustic differences. Another possibility is that the effect is too short-lived to be observed at an ISI of 0 ms. Therefore an identity priming experiment was carried out in which the visual target was presented earlier relative to the prime. The cross-modal version of this task, rather than the intra-modal version, was chosen to ensure that differences in the speed of the lexical decisions would reflect a difference in the degree of lexical activation rather than a difference in the degree of prelexical activation. This argument is in this case even more important since in the identity priming task the phonological overlap between prime and target is considerable, if not complete. This type of overlap can lead to non-lexical facilitation when prime and target are both presented auditorily (e.g., Slowiaczek, McQueen, Soltano & Lynch, 2000). Furthermore, we know from the findings by Spinelli et al. (2003) that the cross-modal identity priming task is sensitive to subtle variation in the initial phoneme.

EXPERIMENT 3.2

Method

Participants

Forty-eight subjects were paid to participate in this experiment. None had participated in the first experiment and none reported any hearing loss.

Materials

The same 40 HF words and 40 LF words of Experiment 3.1 were used in Experiment 3.2, but this time the visual target was the same word as the prime. For each target an unrelated prime was constructed that had the same number of syllables and the same initial phoneme as the related prime. There were also nonword primes. In addition to these items there were 200 filler pairs in which there was no relation between the prime and the target. They consisted of 40 nonword-nonword pairs, 80 nonword-word pairs and 80 word-nonword pairs. All materials came from the same recordings as in Experiment 3.1 and the VOT was manipulated in exactly the same way. The design is summarized in Table 3-1.

Procedure

The procedure was identical to that of Experiment 3.1 except that the visual target was presented 200 ms after the onset of the burst. As in Experiment 3.1, subjects were asked to perform a phoneme identification task on the test items after they had completed the lexical decision task. To shorten the identification phase, listeners only had to identify the initial phoneme of each related prime with the same amount of prevoicing as the one they had heard in the identity priming experiment. Again, half of the items in the identification experiment consisted of distractors (this time words and nonwords) starting with a /p/ or /t/.

Results and Discussion

The results of the identification task indicated that 96% of the items starting with a voiced plosive were identified as voiced. Table 3-2 gives the mean percentage of voiced responses for each prevoicing condition, separately for HF words, LF words and nonwords. The ANOVAs on the transformed proportions showed a significant effect of frequency ($F1(2,94) = 16.51$, $p < .001$; $F2(2,116) = 3.73$, $p < .05$), a significant effect of prevoicing ($F1(2,94) = 95.66$, $p < .001$; $F2(2,232) = 51.19$, $p < .001$) and a significant interaction between frequency and prevoicing ($F1(4,188) = 18.95$, $p < .001$; $F2(4,232) = 6.00$, $p < .001$. Tukey HSD tests showed that the proportion of voiced responses was higher for words (HF or LF) than for nonwords and that, as in Experiment 3.1, tokens without prevoicing were less often identified as voiced than tokens with prevoicing. Furthermore, the interaction between frequency and prevoicing was due to the fact that the difference between tokens with and without prevoicing was larger in the nonwords than in the HF or LF words. This confirms the suggestion which was made earlier, that the identification of voiced plosives without prevoicing was influenced by the lexical status of the item.

Figure 4-3 shows the mean RTs of the lexical decisions for the four priming conditions, plotted separately for the three target conditions. Since correct lexical decisions to the HF and LF targets involved “yes” decisions, while correct lexical decisions to the nonword targets involved “no” decisions, words and nonwords were analyzed separately. In the analysis of the word targets there were signifi-

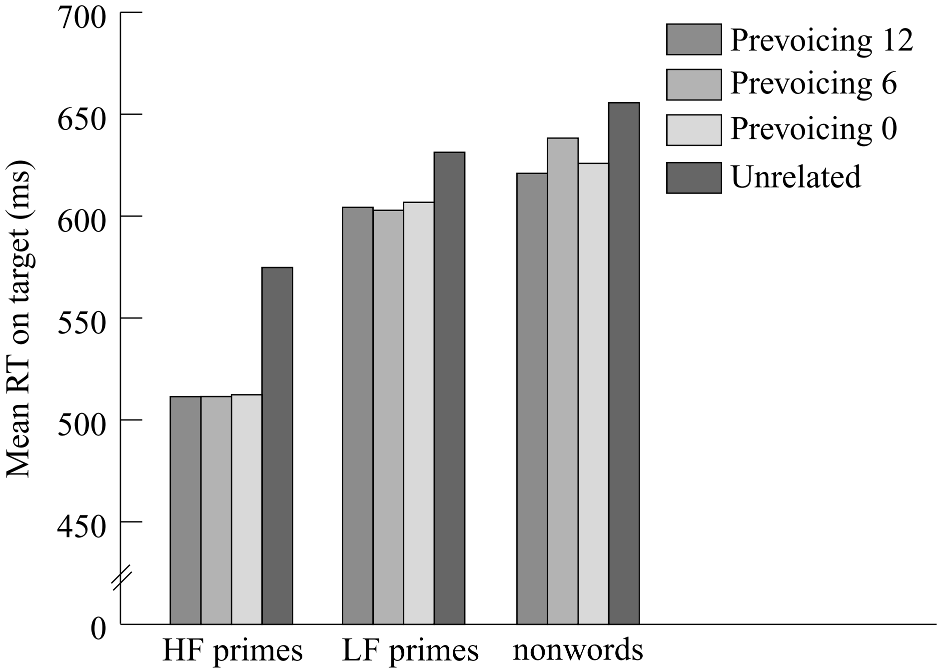

Figure 3-3. *Mean reaction times (RTs) to high frequency (HF) word targets, low frequency (LF) word targets and nonword targets in each of the four priming conditions in Experiment 3.2 (identity priming)*

cant effects of prime type: F1(3,141) = 70.76, p < .001; F2(3,225) = 58.60, p < 001, and frequency: F1(1,47) = 324.26, p < 001; F2(1,75) = 102.41, p < 001. There was also a significant interaction between prime type and frequency: F1(3,141) = 6.50, p < .001; F2(3,225) = 5.38, p = .001).

Only one pairwise comparison was significant: the difference between the prevoicing 6 condition and the unrelated condition: t1(47) = -9.80, p < .001; t2(78) = -8.68, p < .001. This indicates that lexical decisions were significantly faster when targets were preceded by identical primes than when targets were preceded by unrelated primes, and that there was no difference in the degree of priming among the three prevoicing conditions.

The significant effect of frequency indicated that lexical decisions were slower to LF targets than to HF targets (632 ms versus 529 ms). Note that in this experiment related primes and targets were identical and therefore LF primes were followed by LF targets and HF primes by HF targets. The significant interaction between prime type and frequency was further inspected by performing planned t-tests for the three priming condition combinations for LF and HF primes separately. In both frequency groups, only the differences between the prevoicing 6 priming conditions and the unrelated priming conditions were significant. It seems that the interaction reflects the fact that the priming effects were stronger in the HF condition.

Figure 3-3 also shows the mean RTs of the lexical decisions to nonword targets. For the nonwords there was a significant effect of prime type: F1(3,141) = 5.71, p = .001; F2(3,114) = 5.18, p < 0.01. Nevertheless, the three planned t-tests showed that none of the pairwise comparisons were significant, indicating that the priming effect was not substantially present in nonwords. This suggests that the facilitation measured in an identity priming task is mainly due to activation at the lexical level rather than activation at the prelexical level.

The results of the identity priming experiment are comparable to the results of the associative priming experiment. Both experiments show facilitation in lexical decisions to words when targets are preceded by related primes relative to when preceded by unrelated primes. But the VOT manipulations did not affect the degree of facilitation. It is possible, however, that the differences in prevoicing in these experiments were too small to be detectable for the listeners. Even though oscillograms of the stimuli clearly show that they differ in the degree of prevoicing, this does not necessarily mean that listeners can hear these differences. If they cannot hear these differences, it would not be very surprising that this type of variation does not influence lexical access. To test this, a third experiment was conducted in which listeners had to discriminate between the primes with different degrees of prevoicing.

EXPERIMENT 3.3

Method

Participants

Ten subjects participated. None reported any hearing loss and none had taken part in the first two experiments.

Materials

The materials used in this experiment consisted of the 120 test items of Experiment 3.2 (40 HF words, 40 LF words and 40 nonwords). For each item all three prevoicing versions were used (12 periods of prevoicing, 6 periods of prevoicing and no prevoicing). For each item 6 pairs were constructed in such a way that all combinations of prevoicing (prev) appeared: prev 12 - prev 12, prev 6 - prev 6, prev 0 - prev 0 (the “same” pairs) and prev 12 - prev 0, prev 6 - prev 0 and prev 12 - prev 6 (the “different” pairs). The order of items within the “different” pairs was balanced. In total there were 720 pairs.

Procedure

All pairs were presented auditorily in a sound-damped booth in random order. The ISI within a trial was 300 ms and the interval between the offset of a trial and the onset of the next trial was 2000 ms. Subjects were asked to listen carefully to the two items while concentrating on the beginning of the initial sounds and to decide whether the two items were the same or different, by pressing the appropriate button. Before the real experiment started they heard 12 pairs that were different. They had been told beforehand that there was a difference in the initial phoneme between the two items of each pair. After that there was a training session of 24 pairs prior to the main experiment session.

Results and Discussion

Following Macmillan and Creelman (1991), d' was calculated for each subject for each prevoicing combination. Figure 3-4 shows the mean d' values for the three prevoicing combinations.

A one-way ANOVA on the d's indicated that there was a main effect of prevoicing combination: F(2,20) = 24.29, p < .001. A Tukey HSD test showed that the combination prev 12 - prev 0 differed significantly from the combinations prev 6 - prev 0 and prev 12 - prev 6, but that the difference between the combinations prev 6 - prev 0 and prev 12 - prev 6 was not significant. Thus it was easier

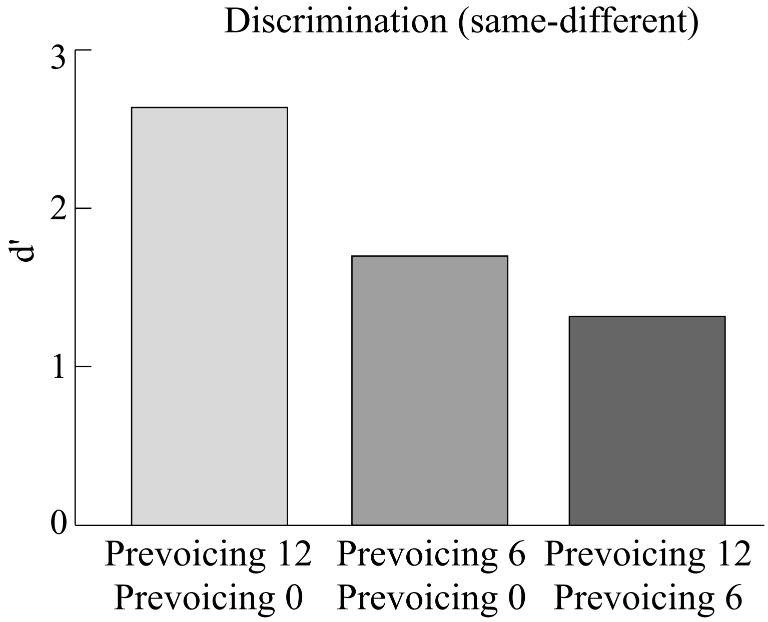

Figure 3-4. *Mean d's for the three combinations of prevoicing in Experiment 3.3 (discrimination)*

to discriminate between two members which differed 12 periods of prevoicing from each other than to discriminate between two members which differed only in 6 periods from each other. Nevertheless, all d's differed significantly from zero (prev 12 - prev 0: $t(10) = 8.20$, $p < .001$; prev 6 - prev 0: $t(10) = 7.11$, $p < .001$; prev 12 - prev 6: $t(10) = 4.34$, $p = .001$). This indicates that listeners performed above chance and could thus discriminate among all three prevoicing durations. The d's of the pairs involving 0 periods of prevoicing also differed significantly from unity (prev 12 - prev 0: $t(10) = 5.10$, $p < .001$; prev 6 - prev 0: $t(10) = 2.92$, $p < .05$). This suggests that the difference between 12 and 6 periods of prevoicing was the most difficult to detect.

THE INFLUENCE OF VOICELESS LEXICAL COMPETITORS ON THE EFFECT OF PREVOICING DIFFERENCES

The two priming experiments reported here investigated the influence of prevoicing variation on lexical access and its interaction with the frequency of the lexical candidates. Both experiments showed a clear priming effect for words. In the associative priming task, listeners were faster to decide that the visual target *roos* (rose) was a word when its was preceded by a semantically related prime such as *bloem* (flower) than when it was preceded by a semantically unrelated prime such as *baan* (job). In the identity priming task, lexical decisions to word targets such as *bloem* were faster when the target was preceded by a prime which was identical to the target (in this case the auditory version of *bloem*) than when the target was preceded by unrelated primes, such as *baan* (job). For the nonwords no substantial priming effect was found. Furthermore, both experiments showed no difference among primes with 12, 6 or 0 periods of prevoicing. Prevoicing variation seemed not to influence lexical access.

The absence of an effect of prevoicing duration (12 periods of prevoicing versus 6 periods of prevoicing) was expected. Recall that van Alphen and Smits (submitted) found that in Dutch the amount of prevoicing appears to be uninformative to the listener. All tokens which were produced with prevoicing were unambiguously identified as voiced, regardless of the exact amount of prevoicing. The primary cue to the perception of [+voice] appeared to be the presence of prevoicing, rather than the duration of prevoicing. Therefore, we predicted that variation in the amount of prevoicing (12 versus 6 periods) would not affect lexical access.

The absence of an effect of the deletion of prevoicing, however, seems puzzling. Van Alphen and Smits showed that when prevoicing was absent, the probability that the token was voiced decreased. Nevertheless, the majority of the voiced plosives without prevoicing were still perceived as voiced. This is in line with the present identification results: although all tokens in these experiments were in general perceived as being voiced, the percentage of voiced responses was lower for tokens with no prevoicing than for tokens with 12 or 6 periods of prevoicing. As mentioned earlier, several studies have shown that lexical activation is sensitive to fine-grained acoustic information, suggesting that information flows continuously from a prelexical level of processing to the lexical level. Based on these findings one would expect that the deletion of prevoicing would affect the degree of activation of lexical candidates starting with voiced plosives.

It is also the case, however, that prevoicing deletion does not result in unnatural or rare tokens. Prevoicing is frequently absent in naturally produced tokens of voiced plosives. As a result, Dutch listeners have often encountered words starting with plosives without prevoicing that should have started with voiced plosives (e.g., hearing *bloem* without prevoicing). Therefore, Dutch listeners might have learned that a plosive without prevoicing could still be voiced. This can explain why no effects of prevoicing deletion were found in the current priming studies. However, since in natural speech most plosives without prevoicing are actually voiceless, listeners should not ignore the presence or absence of prevoicing. Tokens without prevoicing should thus activate both voiced and voiceless prelexical representations, which in turn should activate lexical candidates starting with voiced plosives and lexical candidate starting with voiceless plosives. Note that none of the words in the present study had a voiceless word competitor. That is, for all words, changing the voicing category of the initial voiced plosive from voiced to voiceless resulted in nonwords (e.g., *ploem* is not a Dutch word). Therefore, there were no voiceless lexical candidates which could seriously compete with the voiced word candidates. If it is indeed the case that items starting with voiced plosives without prevoicing activate both voiced and voiceless lexical candidates, one would expect to find effects of prevoicing deletion when primes are used which have a voiceless word candidate that could be activated.

Van Alphen and McQueen (submitted; see Chapter 4) therefore investigated the influence of the competitor environment on the effect of prevoicing variation. They ran two cross-modal identity priming experiments similar to Experiment 3.2 of the present study, but, instead of the frequency conditions, they constructed four different lexical status conditions. The first condition, referred to as the *Blue condition*, contained word primes starting with voiced plosives which had no voiceless word competitor, for example *blauw* (*blauw* means blue and *plauw* is not a word of Dutch). This condition is equivalent to that tested in the present Experiment 3.2. The second condition, referred to as the *Bear condition*, contained word primes starting with voiced plosives with voiceless word competitors, for example *beer* (*beer* means bear and *peer* means pear). The third condition, the *Blem condition,* like the nonword condition in Experiment 3.2, contained nonword primes starting with voiced plosives without voiceless word competitors, for example *blem* (neither *blem* or *plem* is a word of Dutch). The final condition, the Brince condition, contained nonword primes starting with voiced plosives which had a voiceless word competitor, for example *brins* (*brins* is not a word of Dutch and *prins* means prince).

In both experiments five priming conditions were used. In addition to the three prevoicing conditions (prevoicing 12, prevoicing 6, prevoicing 0) and the unrelated priming condition which were also used in the present study, a voiceless priming condition was constructed which contained natural recordings of the voiceless word and nonword counterparts of the voiced primes. This voiceless priming condition (e.g., the prime *peer)* served together with the voiced priming condition (e.g., the prime *beer* with 6 periods of prevoicing) as reference conditions for the condition with voiced primes without prevoicing. The combination of five priming conditions and four lexical status conditions resulted in 20 different conditions in each experiment.

The difference between the two priming experiments was the nature of the target. In the first experiment the items starting with voiced plosives served as targets, while in the second experiment the voiceless counterparts of the voiced items served as targets. For example, in the Bear condition, in the first experiment the target was *beer* and in the second experiment it was *peer.* In this way the degree of activation of both the voiced and voiceless lexical candidates (e.g., *beer* and *peer*) could be measured.

The results showed clear priming effects when prime and target were identical. Furthermore, the RT patterns showed that when prime and target differed only in the voicing of the initial phoneme (for example, the prime was *peer* and the target was *beer*) no facilitatory effect was found. As in the present experiments, there was never an RT difference between the effect of a prime with 12 periods of prevoicing and that of a prime with 6 periods of prevoicing. That is, variation in the amount of prevoicing had no effect in any of the four lexical status conditions. As in the present experiments, there was also no effect of prevoicing deletion when the voiced prime had no voiceless word competitor. Crucially, however, when the voiced prime did have a voiceless word competitor, effects of prevoicing deletion were found. For example, when word targets such as *peer* were preceded by voiced primes without prevoicing (*beer* without prevoicing), lexical decisions to targets were faster in comparison to the same targets preceded by voiced primes with prevoicing (e.g., *beer* with 6 periods of prevoicing), but slower in comparison to the same targets preceded by voiceless primes (e.g., *peer*). Similarly, when a nonword target such as *brins* was preceded by voiced primes without prevoicing (e.g., *brins* without prevoicing), lexical decisions (in this case “no” decisions) were slower in comparison to decisions to the same targets preceded by voiced primes with prevoicing (e.g., *brins* with 6 periods of prevoicing) and faster in comparison to these targets preceded by voiceless primes (e.g., *prins*). For a detailed description of the patterns found in all conditions see van Alphen and McQueen (submitted).

These results suggest that items starting with voiced plosives without prevoicing activate lexical candidates starting with voiced plosives and lexical candidates starting with voiceless plosives. In none of the lexical status conditions was an effect of prevoicing duration (12 versus 6 periods of prevoicing) found. It thus appears that only the acoustic detail which is relevant for the distinction between voiced and voiceless plosives, and thus for the distinction between words starting with voiced plosives and words starting with voiceless plosives (6 versus 0 periods of prevoicing), is passed on to the lexical level. Acoustic detail which is not relevant (12 versus 6 periods of prevoicing) is normalized away at the prelexical level. Effects of the absence of prevoicing, which is relevant for word recognition, were only observed when there was a voiceless lexical candidate. When there was no such candidate, the voiced word candidate was the only plausible lexical hypothesis and could easily win the competition with all other candidates, even when there was no prevoicing.

CONCLUSIONS

In this chapter I have focussed on the phonological voicing distinction in Dutch initial plosives. The phonological distinction between [b] and [d] on the one hand and [p] and [t] on the other hand, is straightforward: the former are voiced and the latter are voiceless. The phonetic realization of this distinction in Dutch, however, is less straightforward. Voiced plosives are said to be produced with voicing during the closure (i.e., with a negative VOT) while voiceless plosives are produced without voicing during the closure but with little or no aspiration (i.e., with a positive VOT). The study on the occurrence of prevoicing in Dutch revealed that a considerable proportion of voiced plosives (25%) were produced without prevoicing. When the aerodynamic circumstances made it more difficult to produce vocal vibration, prevoicing was often absent. Nevertheless, these tokens could still be perceived as voiced, provided that the remaining acoustic cues were in favor of a voiced plosive. This last condition, however, was not always met. As a result, some of the voiced tokens without prevoicing were perceived as voiceless. In contrast, all tokens produced with prevoicing were perceived as voiced.

The presence of prevoicing is thus a very strong cue to the perception of plosives as voiced, but is not fully reliable. This is an intriguing paradox. How does the speech perception system treat the absence of prevoicing? On the one hand, listeners have learned that the absence of prevoicing strongly signals that the token is voiceless; on the other hand, listeners have often encountered words with plosives without prevoicing which appeared to be voiced. The identification results showed that the absence of prevoicing influenced the proportion of voiced responses. Although the majority of the tokens without prevoicing were perceived as voiced, some listeners perceived some of these tokens as voiceless. What consequences does this have for the recognition of words starting with voiced plosives without prevoicing? Is a word like *bloem* still recognized when it is produced without prevoicing, or is it sometimes recognized as *ploem*? Even if it is correctly recognized, the absence of prevoicing could still have affected the recognition process. It is possible that words starting with plosives without prevoicing are more difficult to recognize than words with prevoicing. In order to fully understand the role of prevoicing in perception it is therefore important to also include word recognition.

Two priming experiments were presented which investigated the effects of prevoicing deletion on lexical activation. The difference between the presence and absence of prevoicing was contrasted with a difference in the amount of present prevoicing. Both prevoicing differences were of the same size and fell within the natural range of prevoicing variation. These two types of prevoicing variation differ, however, in their informational value. While the presence or absence or prevoicing is relevant to the voicing distinction, the exact duration of the prevoicing is not. The hypothesis was that only acoustic detail which would help to distinguish between two phoneme classes, and thus between lexical candidates, would affect lexical activation, while irrelevant acoustic detail would be normalized away at the prelexical level. Therefore, the difference involving the presence or absence of prevoicing was expected to affect lexical activation, while the difference in the exact amount of prevoicing was not.

The results suggested that neither of the two differences in prevoicing had an effect on the degree of lexical activation. I argued that the absence of an effect of the deletion of prevoicing could be explained by the fact that prevoicing is frequently absent in Dutch. Dutch listeners have often encountered words starting with plosives without prevoicing that should have started with voiced plosives (e.g., hearing *bloem* without prevoicing). Therefore, they might have learned that a plosive without prevoicing could still be voiced. When the words starting with voiced plosives had no matching voiceless word competitor, as was true for all primes in the two present priming experiments, the lexical candidate starting with the voiced plosive was considered to be the only plausible lexical hypothesis when listeners heard these words without prevoicing. This argument was strengthened by the results of two different priming experiments in which the competitor environment of the words with initial voiced plosives was manipulated (van Alphen & McQueen, submitted). When the primes had a voiceless word competitor, an effect of prevoicing deletion was observed. But there was never an effect of variation in the amount of prevoicing.

The results of these priming experiments show that word recognition is sensitive to phonetic detail that is relevant for lexical distinctions, while irrelevant information is normalized away at an earlier stage of processing. It also shows the robustness of the word recognition system and the influence of the lexical competitor environment. When there is no voiceless word competitor, the recognition system can easily recover from effects of prevoicing deletion, probably due to the fact that this type of variation naturally occurs in Dutch. Only when there is a voiceless word competitor can prevoicing deletion affect the recognition process.

By combining the results of production and perception experiments, including experiments involving word recognition, I aimed to give more insight into the voicing distinction in Dutch initial plosives and the role of prevoicing. In particular, I intended to show that the distinction between voiced and voiceless plosives is less straightforward than one might expect on the basis of the phonological description of these sounds. It appears that there is not a simple binary distinction between voiced and voiceless plosives, but that there are different degrees of voicing. The phonetic realization of a plosive determines the probability that the plosive is voiced. Of all acoustic properties, prevoicing appears to be one of the most important cues affecting this probability. Nevertheless, voiced plosives are frequently produced without prevoicing. This not only affects the role that prevoicing plays in the perception of plosives as voiced or voiceless, but also the way in which the word recognition system treats variation in prevoicing.

THE EFFECT OF VOICE ONSET TIME DIFFERENCES ON LEXICAL ACCESS IN DUTCH

CHAPTER 4

Petra M. van Alphen and James M. McQueen (submitted), *Journal of Experimental Psychology; Human Perception and Performance*

ABSTRACT

Effects on spoken-word recognition of prevoicing differences in Dutch initial voiced plosives were examined. In two cross-modal identity priming experiments, participants heard prime words and nonwords beginning with voiced plosives with 12, 6 or 0 periods of prevoicing, or matched items beginning with voiceless plosives, and made lexical decisions to visual tokens of these items. 6-period primes had the same effect on lexical decisions as 12-period primes. 0-period primes had a different effect, but only when their voiceless counterparts were real words. Listeners could discriminate the 6-period primes from the 12- and 0-period primes. Phonetic detail appears to influence lexical access only when it is useful: In Dutch, presence versus absence of prevoicing is informative; amount of prevoicing is not.

INTRODUCTION

The speech signal contains a wide variety of acoustic properties which encode the words that were intended by the speaker. In order to recognize these words, the listener has to decode the incoming speech signal by extracting phonetic information from the speech signal and mapping this onto stored representations in the mental lexicon. Speech is highly variable, however, and not all acoustic information is relevant for the recognition of words. While some of this information serves to cue particular perceptual distinctions and therefore helps the listener to recognize words, other parts of the signal are not informative in this way and, at least with respect to lexical access, should be regarded as noise. To function optimally, the recognition system must be able to distinguish the relevant acoustic properties from the irrelevant properties in the speech signal. Only the phonetic detail that is relevant for lexical distinctions should therefore affect lexical processing, while irrelevant information should be normalized away at an earlier stage of processing.

We tested this prediction by examining how variation in Voice Onset Time (VOT) influences spoken-word recognition in Dutch. VOT is the primary cue to the phonological distinction between voiced and voiceless plosive consonants in Dutch (van Alphen & Smits, submitted; see Chapter 2). We will argue, however, that not all differences in VOT, even though they fall within the range of natural productions, are informative to listeners. The way in which VOT varies in Dutch gave us the opportunity to examine the effect on lexical access of a difference in VOT values which is critical to the distinction between voiced and voiceless plosives, and the effect of another difference in VOT values which, though quantitatively the same, is not relevant to the voicing distinction. We will argue that the former difference influences lexical access, while the latter does not. Furthermore, we will compare the effects of VOT variation on lexical access in Dutch with similar effects in English (e.g., Andruski, Blumstein & Burton, 1994) and argue that the differences in the effects between the two languages can be explained by differences in the informational value of VOT in the two languages.

Listeners seem to have no problem distinguishing information that is relevant for lexical distinctions from uninformative components of the speech signal. They are capable of understanding words in different phonetic contexts, spoken by a large variety of speakers under many different listening conditions. The human recognition system must therefore be very flexible and be able to extract relevant information from the speech signal in spite of the variability in that signal. In order to be able to function properly, the recognition system must find the right degree of tolerance in dealing with acoustic variation. If the system is too strict, particular phonemic distinctions may not be detected; if the system is too tolerant, particular tokens of segments may be mistakenly assigned to incorrect phonemic categories. As a result, word recognition may be hindered, or the wrong words may be recognized. Research has shown that the recognition system is in fact relatively intolerant of phonemic mismatch between the speech signal and the information stored in lexical representations (e.g., Connine, Blasko & Titone, 1993; Connine, Titone, Deelman & Blasko, 1997; Frauenfelder, Scholten & Content, 2001; Marslen-Wilson, Moss & Halen, 1996; Marslen-Wilson & Zwitserlood, 1989; Milberg, Blumstein & Dworetzky, 1988). Mismatch in the initial phoneme of the word can inhibit lexical access (Marslen-Wilson & Zwitserlood, 1989); a mismatch later in the word can produce rapid deactivation of lexical candidates (Frauenfelder, Scholten & Content, 2001). Phonemic mismatches can thus cause large changes at the lexical level. In fact, there is a growing body of evidence which suggests that the lexical level is sensitive to very fine-grained differences in the speech signal.

The results of Connine et al. (1993, 1997) suggest that lexical access is not only sensitive to phonemic mismatch, but also to the number of mismatching features causing the phonemic mismatch: a word like *tulip* is activated more strongly upon hearing *dulip* (the [d] differs from the [t] only in voicing) than upon hearing *vulip* (the [v] differs from the [t] in voicing, place and manner of articulation). The claim that the mapping between the speech input and stored lexical representations is sensitive to subphonemic variation is supported by a number of other studies investigating various types of subphonemic variation (see McQueen, in press, for a detailed overview). Lexical access appears to be sensitive to mismatching formant transitions resulting from cross-splicing vowels and their following consonants (e.g., Dahan, Magnuson, Tanenhaus & Hogan, 2001; Marslen-Wilson & Warren, 1994; McQueen, Norris & Cutler, 1999; Streeter & Nigro, 1979; Whalen, 1984, 1991), to fine-grained acoustic information about syllable and word boundaries (Davis, Marslen-Wilson & Gaskell, 2002; Gow & Gordon, 1995; Salverda, Dahan & McQueen, 2003; Spinelli, McQueen & Cutler, 2003; Tabossi, Collina, Mazzetti & Zoppello, 2000), and to subphonemic cues to assimilation of place of articulation (Gow, 2002).

Further evidence that lexical activation is influenced by subphonemic variation comes from studies investigating a type of subphonemic variation that is of particular interest for the present study, namely variation in VOT. VOT is the interval between the release of the occlusion for a plosive consonant and the moment that the vocal folds start vibrating. In English, voiced plosives are produced with a slightly positive VOT, while voiceless plosives are produced with a longer positive VOT. English listeners are very sensitive to VOT differences since VOT is one of the major cues for the voicing distinction in English (e.g., Lisker & Abramson, 1970; Miller & Volaitis, 1989).

Andruski et al. (1994) examined the influence on lexical access in English of VOT variation within the voiceless category of initial plosives. Half of the words in their experiment had a voiced word competitor, that is, changing the initial voiceless plosive into the matching voiced plosive resulted in a word, for example *pear* (*bear)*. The other half of the words did not have a lexical competitor, for example *king (ging)*. Three VOT conditions were created by either retaining the original VOT of the initial voiceless plosives, or by removing one third of the original VOT, or by removing two thirds of the original VOT. An identification study indicated that all three VOT variations were in general perceived as voiceless. The words in the three VOT conditions served as related primes in an associative priming task in which both prime and target were presented auditorily. Listeners were asked to decide whether the target (such as *queen* after the prime *king*) was a word or a nonword. Listeners were faster to make lexical decisions to targets when these were preceded by the related primes, starting with various VOTs, than when they were preceded by unrelated primes. Critically, when the interstimulus interval between prime and target was 50 ms, lexical decisions to targets preceded by primes of which two thirds of the VOT was removed were significantly slower than lexical decisions to the same targets preceded by primes with unaltered VOT. Although lexical decisions to targets preceded by primes with voiced word competitors were overall slower than to targets preceded by primes without voiced word competitors, the competitor environment seemed not to influence the effects of VOT variation. The results of Andruski et al. indicate that, in English, variations in the positive VOTs of initial voiceless plosives affect the degree of activation of lexical candidates starting with voiceless plosives.

Similar effects have been observed by Utman, Blumstein and Burton (2000), using the identity priming task. Spoken target words and nonwords were preceded by the same natural tokens of those words and nonwords, or by tokens in which the VOT was shortened. When targets and primes were words, such as *kiss*, primes in which the VOT was shortened resulted in reduced priming effects relative to primes which were identical to the targets. In contrast, when targets and primes were nonwords, no effect of the VOT reduction was found on the lexical decisions. McMurray, Tanenhaus and Aslin (2002) also investigated the effect of VOT variation on lexical access in English, using the eye-tracking paradigm. They found that the mean proportion of fixations to one of two target pictures which displayed the two members of a minimal word pair like *bear-pear* varied gradually as a function of the VOT of the initial plosives.

All these studies show that lexical access is influenced by subphonemic variation and support current models of word recognition such as TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994; Norris, McQueen & Cutler, 2000), in which information flows continuously from a prelexical level of processing to the lexical level. At the prelexical level, the incoming speech signal is normalized and useful information is extracted from the speech signal and translated into abstract representations. There are, unfortunately, no experimental data which unambiguously inform us about the size of the representational units at the prelexical level. Many different units have been proposed, including syllables (Mehler, 1981), semi-syllables (Massaro, 1987), phonemes (Foss & Blank, 1980; Nearey, 2001), allophones (Luce, Goldinger, Auer & Vitevitch, 2000) and features (Stevens, 2002). Although we do not want to make any claims about the size of these prelexical units, we do assume that prelexical units exist, that they are activated in proportion to their match with the speech signal, and that they pass their activation continuously up to word representations at the lexical level. Another assumption we make which is important for the interpretation of the results of the experiments we will present is that the lexical candidates that are activated as a result of the acoustic input will compete with each other for recognition. The candidate that matches the acoustic input best will eventually suppress the activation of the other candidates. The competition process acts to make the recognition process more efficient. There is a large amount of evidence from different experiments using different experimental tasks for the existence of competition between lexical candidates (see McQueen, in press, for an overview). All current models of spoken word recognition therefore assume that there is some form of lexical competition.

Given the evidence that the prelexical processor does not act in a serial and categorical way, but passes activation continuously on to the lexicon, the question one then has to ask is whether the prelexical processor passes on all fine-grained phonetic information to the lexicon, or whether particular types of variation are normalized away at the prelexical level of processing. As described earlier, the speech signal is highly variable and part of this variation is not informative, that is, it does not help the listener to understand the speaker’s message. If the activation of prelexical units were to be influenced by any type of acoustic variation, and this activation were then passed on to the lexicon, the recognition of words could be hindered by irrelevant acoustic variation. We therefore hypothesize that the prelexical processor only passes on information that has value in communication, that is, information that helps listeners to distinguish between phonemes and thus between words. The aim of the present paper was to test this hypothesis by measuring word recognition while varying VOT in Dutch initial voiced plosives ([b] and [d]).

In Dutch, the main perceptual cue for the distinction between voiced and voiceless initial plosives appears to be the presence or absence of prevoicing (van Alphen & Smits, submitted). Prevoicing is the presence of vocal fold vibration during the consonantal closure, and corresponds to a negative VOT since the moment that the vocal folds start vibrating falls before the moment of the release of the occlusion. One of the requirements for the production of vocal fold vibration is that sufficient transglottal pressure is obtained and maintained to let the vocal folds vibrate. This is relatively hard when all outgoing air pathways are closed, as is the case with plosive consonants. Given the effort it takes to produce prevoicing, it is very likely both that it is difficult for speakers to control the exact duration of prevoicing and that they may sometimes fail to produce any prevoicing. Van Alphen and Smits indeed found that the duration of prevoicing varies considerably within and among speakers and that no less than 25% of the voiced tokens in their sample were produced without prevoicing. Their perception data showed that when a plosive was produced with prevoicing, listeners perceived the plosive unambiguously as being voiced. This was true for both [b] and [d]; note that the velar plosive [g] is marginal in Dutch, since it only occurs in loan words. The velar plosive was, therefore, not tested by van Alphen and Smits, and will not be examined here. Furthermore, van Alphen and Smits observed that the exact duration of the prevoicing did not influence the voicing percept, as long as there was prevoicing present. When listeners had to identify the voiced plosives which were produced without prevoicing, however, the majority of the unprevoiced tokens were still perceived as being voiced, as a result of other cues in the signal. Perception of labial plosives without prevoicing as voiced depended on the F0 movement immediately after the burst, while perception of alveolar plosives without prevoicing as voiced depended on the spectral centre of gravity of the burst.

In summary, the duration of prevoicing of Dutch initial plosives varies considerably and, in line with this, the exact duration of prevoicing seems not to affect the strength of the evidence that the plosives are voiced. Thus, the exact duration of prevoicing does not have any informational value. The information that helps the listeners to distinguish between voiced and voiceless plosives lies in the presence or absence of prevoicing. When prevoicing is present, the plosive is unambiguously perceived as voiced. Deleting the prevoicing of voiced plosives therefore reduces the probability that the plosive is voiced (although the plosive will still be perceived as voiced if the remaining cues support this conclusion). This difference in the informational value of the exact duration of prevoicing and the presence or absence of prevoicing gives us the opportunity to test the hypothesis that only phonetic detail that is informative will be passed on to the lexical level.

The present research project examined two prevoicing differences: the difference between 12 and 6 periods of prevoicing (corresponding to approximately 136 and 68 ms of prevoicing), and the difference between 6 and 0 periods of prevoicing (corresponding to 68 and 0 ms of prevoicing). Both differences are of the same size, namely 6 periods of prevoicing, and both differences vary along the VOT scale within the natural range of VOT variation. The difference between the two is that the difference between 12 and 6 periods of prevoicing involves only a difference in the duration of the prevoicing, while the difference between 6 and 0 periods also involves the difference between the presence and absence of prevoicing. The findings of van Alphen and Smits (submitted) suggest that the difference between 12 and 6 periods of prevoicing would not affect the probability that the plosive is voiced, since in both cases there is prevoicing present. In contrast, the difference between 6 and 0 periods of prevoicing would indeed affect that probability, as deleting the prevoicing entirely takes away a major cue to the voicing of the plosive. We therefore predicted that the difference between 12 and 6 periods of prevoicing would be normalized away at the prelexical level and would thus not affect lexical access, but that the difference between 6 and 0 periods of prevoicing would result in different degrees of activation of the prelexical representations, resulting in different degrees of activation at the lexical level.

Interestingly, Dutch voiced plosives are not always produced with prevoicing. Thus, the absence of prevoicing in Dutch occurs both in voiced and voiceless plosives. Therefore, it is possible that plosives without prevoicing would not only result in weaker activation of the voiced prelexical representation (e.g., the representation for [b]) in comparison to voiced plosives with prevoicing, but that the plosive without prevoicing would also activate the voiceless prelexical representation to some extent (e.g., the representation for [p]). If this is true (assuming that the prelexical processor passes information continuously on to the lexical level), then one would predict that upon hearing words starting with voiced plosives without prevoicing not only lexical candidates starting with voiced plosives, but also lexical candidates starting with voiceless plosives would be activated.

In order to test this, the present study investigated the two types of prevoicing variation (12 versus 6 and 6 versus 0 periods of prevoicing) in plosives at the beginning of both words and nonwords which had either a voiceless word competitor or not. English examples of these materials are as follows: the words *blue* (where *plue* is a nonword) and *bear* (with the voiceless lexical competitor *pear*); and the nonwords *blem* (*plem* is also a nonword) and *brince* (with the lexical competitor *prince*). Two cross-modal identity priming experiments and a discrimination experiment were carried out using Dutch versions of these four types of item (see Table 4-1 for examples). Experiment 4.1 examined the influence of spoken primes varying in prevoicing on recognition of visual targets beginning with voiced plosives (i.e., Dutch versions of *blue, bear, blem* and *brince*). Experiment 4.2 used the same primes, but measured recognition of the voiceless counterparts of the Experiment 4.1 targets (i.e., Dutch versions of *plue, pear, plem* and *prince*). Experiment 4.3 tested listeners’ ability to discriminate between the primes used in the first two experiments.

Table 4-1. *Design of Experiments 4.1 and 4.2. For each combination of priming condition and lexical status condition examples of a prime and target are given. Real words have their English translation in parentheses.*

		Lexical Status Condition			
	Priming Condition	Blue condition	Bear condition	Blem condition	Brince condition
PRIME	Prevoicing 12	blauw *(blue)*	beer *(bear)*	blem -	brins -
	Prevoicing 6	blauw *(blue)*	beer *(bear)*	blem -	brins -
	Prevoicing 0	blauw *(blue)*	beer *(bear)*	blem -	brins -
	Voiceless	plauw -	peer *(pear)*	plem -	prins *(prince)*
	Unrelated	buurt *(neighborhood)*	breuk *(fraction)*	burf -	bleug -
TARGET	Voiced (Experiment 4.1)	blauw *(blue)*	beer *(bear)*	blem	brins
	Voiceless (Experiment 4.2)	plauw	peer *(pear)*	plem	prins *(prince)*

EXPERIMENT 4.1

The purpose of Experiment 4.1 was to investigate the effect on lexical activation of prevoicing variation in Dutch initial plosives. Three prevoicing durations were chosen such that the smallest duration was zero and such that the physical difference between the subsequent prevoicing durations was the same: 0, 6 and 12 periods of prevoicing. Importantly, all three variations of prevoicing fell within the natural range of prevoicing duration found by van Alphen and Smits (submitted). On the basis of van Alphen and Smits we predicted that plosives with 12 and 6 periods of prevoicing would both be perceived as voiced plosives, while deleting the prevoicing would result in less clear voiced plosives. Nevertheless, we predicted that these unprevoiced plosives would in general still be perceived as being voiced. These predictions were tested in the identification phase of the experiment.

In the main phase of the experiment, the cross-modal identity priming task was used to explore the influence of prevoicing variation on lexical activation. Within-modality identity priming has been used in previous studies on VOT variation (Utman et al., 2000). The underlying idea in the use of this task is that the processing of a stimulus (the prime) may facilitate or inhibit the subsequent processing of a following stimulus (the target) if the phonological overlap between the stimuli coincides with units that are involved in word recognition. To measure the influence of the presentation of the prime on the processing of the target, participants are asked to perform a task such as lexical decision on the targets. The reaction times (RTs) of these decisions are then compared to the RTs in a baseline condition, in which the target is preceded by a phonologically unrelated prime. When primes and targets are identical, robust facilitation is observed (Slowiaczek & Hamburger, 1992; Slowiaczek & Pisoni, 1986). Many studies have also shown, however, that when primes and targets rhyme, for example *list* and *fist*, lexical decisions on the targets are faster than when the targets are preceded by phonologically unrelated primes (Norris, McQueen & Cutler, 2002; Praamstra, Meyer & Levelt, 1994; Radeau, Besson, Fontaneau & Castro, 1998; Radeau, Morais & Segui, 1995; Slowiaczek, McQueen, Soltano & Lynch, 2000). In all these studies both primes and targets were presented auditorily. When primes are spoken and targets are presented visually, however, the presentation of rhyming primes does not produce facilitation (Cutler, van Ooijen & Norris, 1999; Radeau, Segui & Morais, 1994). These findings suggest that within-modality identity priming (e.g., in Utman et al., 2000) may in part reflect prelexical effects (i.e., effects due to the overlap of sublexical components of primes and targets rather than lexical overlap). We therefore avoided this version of the priming paradigm.

We chose the cross-modal version instead. Cross-modal priming experiments have shown that when primes are phonologically identical to the targets, a significant facilitatory effect is observed (Norris, Cutler, McQueen & Butterfield, submitted; Spinelli, McQueen & Cutler, 2003). But, as we have just pointed out, no such facilitatory effects are found when prime and target differ only in their initial sounds (Cutler, van Ooijen & Norris, 1999; Radeau, Segui & Morais, 1994). This suggests that the facilitation measured in the cross-modal identity priming task is mainly due to activation at the lexical level rather than activation at the prelexical level. Since the task appears to be sensitive to subtle variation in the initial phoneme (Spinelli et al., 2003), we hoped it would be sensitive to differences in prevoicing in initial plosives.

In Experiment 4.1 participants were presented with target words and nonwords starting with voiced plosives (e.g., the words *blauw* (blue), and *beer* (bear), and the nonwords *blem* and *brins*; hereafter, we will refer to these as the *voiced targets*; see Table 4-1). These were preceded by identical spoken primes starting with voiced plosives with different prevoicing durations (the *voiced primes*, with 12, 6 or 0 periods of prevoicing). In addition to these three prevoicing conditions, there was also a priming condition in which primes were identical to the targets, except for the initial plosive which was voiceless instead of voiced (hereafter, the *voiceless primes*), and a priming condition in which primes were phonologically unrelated to the targets (the *unrelated* primes).

On the basis of the findings in previous studies using the cross-modal identity priming paradigm, as described above, we expected to find the following patterns. First, we expected to find a facilitatory effect when the primes were identical to the targets (that is, in the conditions where the voiced targets were preceded by voiced primes with prevoicing) in comparison to unrelated primes. As described earlier, we expected that there would be no difference between the effect of primes with 12 periods of prevoicing and that of primes with 6 periods of prevoicing, since the exact duration of prevoicing is not relevant for the perception of voiced plosives, as long as there is prevoicing. Therefore, we did not expect to find a difference between the RTs of the lexical decisions to targets preceded by voiced primes starting with 12 periods of prevoicing or with 6 periods of prevoicing. If there were a graded effect of the duration of prevoicing on lexical activation, however, the voiced prelexical representations (e.g., the representations for [b] and [d]) would be activated more strongly upon hearing a voiced plosive with 12 periods of prevoicing than upon hearing a voiced plosive with 6 periods of prevoicing, since the former is further away from the phoneme boundary. One would then expect to find faster RTs to targets preceded by voiced primes with 12 periods of prevoicing than to targets preceded by voiced primes with 6 periods of prevoicing.

Second, we predicted that we would find no facilitatory effect when primes and targets differed only in the initial phoneme (that is, in the conditions where the voiced targets were preceded by the voiceless primes) in relation to the unrelated priming condition. It was necessary to find these two general patterns (facilitatory identity priming and no priming from the voiceless primes), since they could then be used as reference conditions for the third prevoicing condition, namely that with voiced primes without prevoicing. If there is no graded activation of the voiced word candidate as a function of whether prevoicing is present or absent, we should expect to find a pattern similar to the pattern found for voiced primes with prevoicing. If activation of the voiced word candidate does vary as a function of the presence or absence of prevoicing, however, we should expect to find a pattern for the voiced primes without prevoicing which falls in between the two reference conditions (voiced primes with prevoicing and voiceless primes). To test all these effects, five pairwise comparisons were planned. The differences in RTs and errors to targets preceded by the following pairs of primes were tested: voiced primes with 12 periods of prevoicing and voiced primes with 6 periods of prevoicing; voiced primes with 6 periods of prevoicing and unrelated primes; voiceless primes and unrelated primes; voiced primes without prevoicing and voiced primes with 6 periods of prevoicing; and voiced primes without prevoicing and voiceless primes.

We tested both word and nonword targets. The RTs of both correct “yes” responses and correct “no” responses were therefore analyzed. It is important to note that these two responses are a result of two different kind of decisions, namely the decision that a string of graphemes is a word or the decision that it is not a word. The yes/no lexical decision task can be considered to involve two criterion levels of evidence (e.g., Balota & Chumbley, 1984; Stone & Van Orden, 1993; see also Grainger & Jacobs, 1996; Perea, Rosa & Gomez, 2002), namely, a criterion level of evidence for a word response (a positive criterion) and a criterion level of evidence for a nonword response (a negative criterion). When the accumulated evidence given a visual stimulus reaches the positive criterion, a “yes” response will be made, when it reaches the negative criterion, a “no” response will be made. If the word target is preceded by an identical prime, the evidence for a word response will accumulate faster than when the target is preceded by an unrelated prime, since in the former case the phonological representation of the word will also be activated and will spread activation to the visual representation of the word. The visual representation thus receives activation as a result of the presentation of the target and, via the auditory system, the prime. Therefore the word criterion is reached earlier. When the target is a nonword, but is preceded by a word prime that is phonologically very similar to the target, there is a conflict between the evidence given by the auditory input and the visual input. The word prime will activate the phonological word representation which will in turn activate its orthographic representation. As a result of the activation in the visual system, the negative criterion is reached later than when the nonword target is preceded by an unrelated prime, resulting is slower “no” responses in the former case. Therefore, not only does the difference in RTs to word targets preceded by related and unrelated primes reflect the degree of activation of phonological lexical representations given a particular prime, but so also does the difference in RTs to nonword targets preceded by related and unrelated primes.

The use of both words and nonwords beginning with voiced stops allowed us to examine lexical involvement on any effects of prevoicing variation. In addition, we investigated the influence of word competitors starting with voiceless plosives. As mentioned before, removing the prevoicing of voiced plosives takes away an important cue that the plosive is voiced. These unprevoiced plosives are likely to contain enough remaining cues for the plosive to still be perceived as voiced (van Alphen & Smits, submitted), but the absence of prevoicing will shift the voiced plosive closer to the voiceless category in perceptual space. Therefore it is very likely that upon hearing items starting with voiced plosives without prevoicing, any word candidates starting with the voiceless counterpart will also be activated. To test this, targets (and thus voiced primes) were either words or nonwords with or without voiceless word competitors, that is, changing the initial voiced plosive into its voiceless counterpart resulted in either a word or a nonword. This resulted in four different lexical status conditions (see Table 4-1). If it is indeed the case that voiceless word competitors are activated upon hearing voiced primes without prevoicing, we expected to find a difference between the conditions with voiceless word competitors from those without such competitors. We expected that the effect of the voiceless word competitor would be strongest in the condition in which the voiced target was a nonword, since the voiceless word would then be the only strongly activated candidate. In the condition in which both the voiced target and the voiceless competitor were words, the two candidates were expected to compete with each other for recognition. In that case we expected the voiced word candidate to win, since we predicted that voiced plosives without prevoicing would in general still be perceived as voiced. If the voiceless word competitor is not activated upon hearing primes starting with these plosives, no differences should be found as a function of the lexical status of the counterpart.

Method

Participants

Sixty volunteers from the Max Planck Institute participant pool were paid to take part in this experiment. All were native speakers of Dutch, and none reported any hearing loss.

Materials

Forty pairs of Dutch monosyllables in each of four lexical status conditions were selected which consisted of an item starting with a voiced plosive and a matched item that was identical except that its initial plosive was voiceless. These pairs are listed in the Appendix 4-A. The first condition, which we will refer to as the *Blue condition*, contained pairs for which the voiced members were words and the voiceless members were nonwords, for example *blauw-plauw* (*blauw* means blue and *plauw* is not a word of Dutch). The second condition, which we will refer to as the *Bear condition*, contained pairs for which both the voiced and voiceless members were words, for example *beer-peer* (bear-pear). The third condition, the *Blem condition*, contained pairs for which both the voiced members and the voiceless members were nonwords, for example *blem-plem*. Finally, in the *Brince condition*, the voiced members were nonwords and the voiceless members were words, for example *brins-prins* (*brins* is not a word in Dutch and *prins* means prince). Each of the four conditions consisted of 27 pairs starting with a /b/ and a /p/, and 13 pairs starting with a /d/ and a /t/. This ratio of labial and alveolar plosives was determined by the number of items which could be found in the Bear condition. The mean frequency of the voiced words of the Blue pairs was 35 per million words (from the CELEX lexical database, Baayen, Piepenbrock & Gulikers, 1995). The voiceless members of the Brince pairs had a mean frequency of 36 per million words. The frequency of the word members was thus matched across the Blue and Brince conditions. The mean frequency of the voiced word members and the voiceless word members of the Bear pairs was respectively 14 and 15 per million words. So the words of these pairs were also matched in frequency.

There were five priming conditions: prevoicing 12, prevoicing 6, prevoicing 0, voiceless and unrelated. In the first three conditions the primes consisted of the voiced members of the pairs. For each target these three different primes were acoustically identical, except for the duration of the prevoicing, which varied systematically among the three conditions (respectively 12 periods, 6 periods and no periods of prevoicing). In the voiceless priming condition, the primes consisted of the voiceless members of the pairs. In the unrelated priming condition, primes were unrelated to the targets, but started with the same voiced plosives (labial or alveolar) as the targets. The lexical status of each unrelated prime was identical to the lexical status of the voiced prime for that target, but the unrelated primes had no voiceless word competitors.

In addition, there were 320 unrelated prime-target pairs which served as fillers. These pairs consisted of 40 nonword prime - nonword target pairs, 120 nonword prime - word target pairs, 120 word prime - nonword target pairs and 40 word prime - word target pairs. The fillers consisted of both mono- and polysyllabic items. Table 4-1 gives an overview of the four lexical status conditions and the five priming conditions. For each of the lexical status conditions one target example is given, together with its five different primes.

Stimulus construction

All items and fillers were recorded several times by a male native speaker of Dutch in a sound-attenuated booth onto Digital Audio Tape (sampling at 48 kHz with 16-bit resolution). The utterances were redigitized at a sample rate of 16 kHz through the ESPS speech editing system with Xwaves. For the three prevoicing priming conditions and the unrelated condition, tokens were chosen which were produced clearly and with prevoicing. Subsequently, the original prevoicing of each related prevoiced item was replaced by 12, 6 or 0 periods of prevoicing (corresponding to 136, 68 or 0 ms of prevoicing for /b/ and to 138, 69 and 0 ms of prevoicing for /d/), in order to create the three different prevoicing conditions. The first full period of prevoicing plus the lead-in (of 7 ms) of a natural token of the word /bʉs/ (bus) was chosen as the first period of prevoicing for the two conditions with prevoicing (12 and 6 periods) for the items starting with a labial plosive. Similarly, the last prevoicing period of that same token of /bʉs/ always served as the last prevoicing period in these two conditions. The intervening prevoicing periods (10 and 4) were randomly chosen from the /bʉs/ token. The same procedure was applied to create the prevoicing 12 and prevoicing 6 conditions for the items starting with an alveolar plosive, but now the prevoicing periods were derived from a natural token of the word /dʉs/ (thus). To control for any splicing effects, the prevoicing of each of the unrelated primes was also replaced by six periods of prevoicing. The voiceless primes were natural productions of the voiceless counterparts of the targets.

Procedure

The experiment had two parts. The first part consisted of the cross-modal identity priming task. Primes were presented binaurally over headphones at a comfortable listening level. The target was presented visually 200 ms after the onset of the plosive burst of the prime, such that the target appeared on the screen during the prime. In this way, listeners had always heard the prevoicing and burst of the initial plosive before the target appeared. Previous research has suggested that presenting the target in a cross-modal identity priming task at the onset of the prime might be too early to measure differential lexical activation across conditions, and presenting the target at the offset of the prime may be too late (van der Lugt, 1999). Presenting the target halfway through the prime results in differential priming effects (Spinelli et al., 2003). The targets in the present experiment were therefore presented during the primes.

Targets were displayed in lower case Arial 36-point typeface. Listeners were tested in sound-attenuated booths. They were instructed to listen to the auditory stimulus and decide as quickly as possible whether the stimulus on the screen was a word or a nonword by pressing one of two appropriately labeled buttons. Five different lists were constructed in which priming condition was counterbalanced across lists. Each participant thus saw each target only once, preceded by one of the five possible primes for that item. Each list also contained all 320 fillers, such that there was an equal number of word and nonword targets preceded by either a word or nonword prime. Of the total of 480 pairs in a given list, 128 pairs (27%) were related (that is, the prime was identical to the target or the prime differed from the target only in the voicing of the first phoneme). There were two randomized versions of each list.

The second part of the experiment consisted of a phoneme identification task. Each participant heard those prevoicing primes that he or she had heard in the priming experiment (24 per lexical status condition). In addition to the voiced primes, they heard the voiceless counterparts of all primes, such that half of the items in this part of the experiment started with voiced plosives and half with voiceless plosives. In total there were 192 items per listener. Items were presented blocked by place of articulation. Half of the participants started with the labial plosives and the other half with the alveolar plosives. They were instructed to label the first sound of each item as /b/ or /p/, or as /d/ or /t/, by pressing one of two buttons, which were labeled appropriately for that particular block.

Results

First we looked at the distribution of correct responses to the targets in the lexical decision phase. Some of the word targets were of low frequency of occurrence and were often not recognized as words. Furthermore, some of the voiceless word competitors were also of low frequency and might therefore not be regarded as lexical competitors. This could have weakened a possible effect of the lexical status condition (i.e., both the effect of the lexical status of the target and the lexical status of the voiceless counterpart). In order to find out which items and which voiceless counterparts did not have a very clear lexical status, the mean proportion of correct responses was calculated for each of the targets in both Experiment 4.1 and Experiment 4.2 (remember that the targets of Experiment 4.2 were the voiceless counterparts of the targets in Experiment 4.1). To avoid the influence of any priming effects, only the responses to targets in the unrelated priming condition were included in this analysis. All targets for which the proportions of correct responses were more than two standard deviations away from the mean proportion of correct responses (resulting in a cut-off value of 23% errors) were removed from all subsequent analyses in all three experiments. Targets with voiceless counterparts that were misclassified in more than 23% of the cases in Experiment 4.2 were also removed from all further analyses in all three experiments. In total, nine Blue targets, ten Bear targets, six Blem targets and nine Brince targets were excluded (see Appendix 4-A). As a result, the mean frequency of the words in all conditions increased, but was still matched across the Blue and Brince conditions, and across the voiced and voiceless word members of the Bear condition. The mean frequency of the voiced words of the Blue pairs was now 39 per million words (based on the CELEX database) and the voiceless members of the Brince pairs had a mean frequency of 40 per million words. The mean frequencies of the voiced and voiceless words in the Bear pairs were now respectively 16 and 19 per million words.

Identification

We then analyzed the identification responses from the second phase of the experiment. For each combination of prevoicing (prevoicing 12, prevoicing 6 and prevoicing 0) and lexical status (Blue, Bear, Blem and Brince), the percentage of voiced responses were calculated separately for each listener. The mean percentages of voiced responses for all listeners are presented in Table 4-2. These were close to 100% in all lexical status conditions for the prevoicing 12 and prevoicing 6 conditions. Items without prevoicing were identified as being voiced less often than those in the two other prevoicing conditions, but were still more often identified as being voiced than voiceless. The overall correct identification performance on the voiceless items was 93%.

The proportions of voiced responses were converted through an arcsine transformation (Studebaker, 1985) and submitted to repeated measures analyses of variance (ANOVAs) with the factors prevoicing and lexical status. The main effect of lexical status was only significant in the item analysis (F2(3,122) = 4.72, $p < .01$). There was a main effect of prevoicing in both the participants (F1) and items (F2) analyses (F1(2,88) = 160.99, $p < .001$; F2(2,244) = 110.90, $p < .001$). Tukey honestly significant difference (HSD) tests showed that in all four lexical status conditions the proportions of voiced responses to items starting with a plosive without prevoicing were significantly

Table 4-2. *Percentage voiced responses in the identification task of Experiment 4.1*

	Blue condition	Bear condition	Blem condition	Brince condition
Prevoicing 12	98.7	97.9	98.9	97.1
Prevoicing 6	97.3	98.1	98.3	97.9
Prevoicing 0	88.3	80.2	84.6	68.1

smaller than to items starting with plosives with 12 or 6 periods of prevoicing (at the .05 level). Although it seems that this difference was larger in the Brince condition than in any of the other lexical status conditions, the interaction between voicing and lexical status was only significant in the item analysis (F2(6,244) = 4.19, p < .001).

These data show that plosives from which the prevoicing was removed were in general still perceived as being voiced. In the Blem condition (which is the most appropriate condition to consider, since there are no lexical factors present in this condition that might have influenced identification) the proportion of voiced responses for the prevoicing 0 condition was almost 85%. As predicted on the basis of van Alphen and Smits (submitted), removing the prevoicing of a voiced plosive reduces the probability of the plosive being voiced to some extent, but the other acoustic cues to voicing which are still present provide enough evidence for the plosive usually to be identified as voiced.

Lexical Decision

The mean RTs, measured from target onset, of the correct responses to targets in each priming condition are plotted separately for each of the four lexical status conditions in Figures 4-1 and 4-2. Figure 4-1 shows the mean RTs of “yes” responses to word targets (Blue and Bear conditions) and Figure 4-2 shows the mean RTs of “no” responses to nonword targets (Blem and Brince conditions).

We will first focus on the conditions with word targets (Blue and Bear). Figure 4-1 shows a similar pattern for words without voiceless word competitors and words with voiceless word competitors. Repeated measure ANOVAs on these data showed that there was a significant effect of priming condition (prevoicing 12, prevoicing 6, prevoicing 0, voiceless and unrelated): F1(4,236) = 34.90, p < .001; F2(4,236) = 22.04, p < .001, but no effect of competitor (plus or minus lexical competitor) and no interaction. Individual ANOVAs showed that in both lexical status conditions there was main effect of priming:

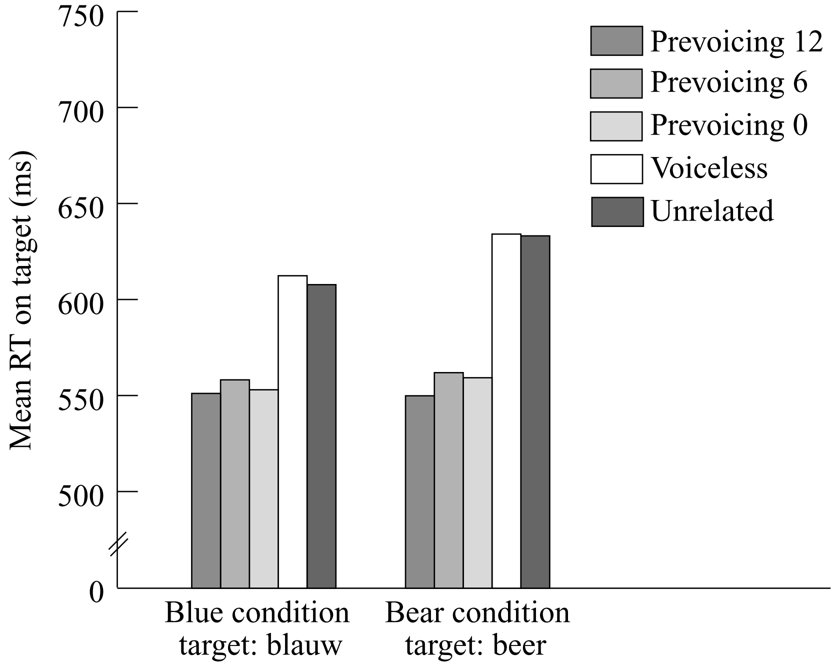

Figure 4-1. *Mean RTs to voiced word targets in the lexical decision task of Experiment 4.1*

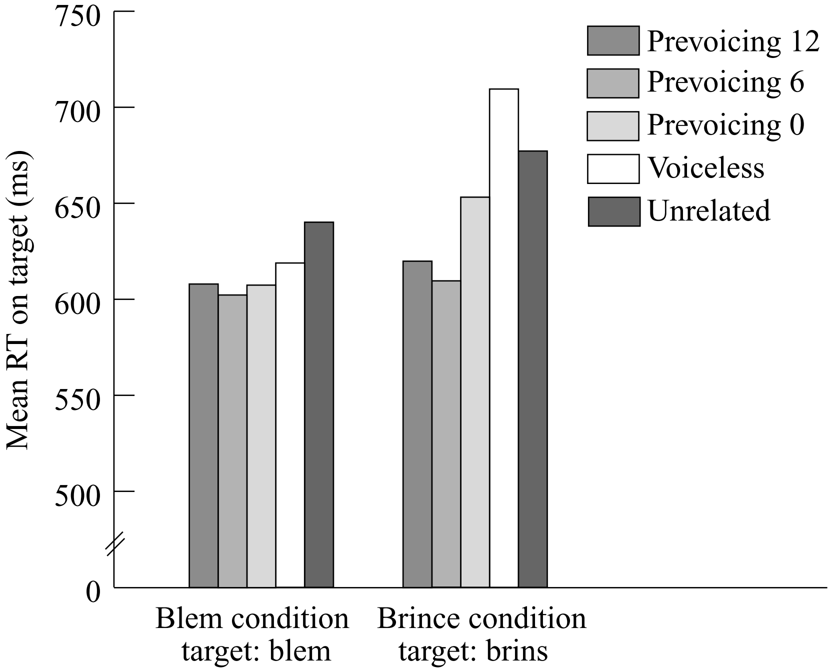

Figure 4-2. *Mean RTs to voiced nonword targets in the lexical decision task of Experiment 4.1*

Table 4-3. *T-tests for planned comparisons in the lexical decision task for each of the five comparisons in each of the four lexical status conditions of Experiment 4.1*

Comparison	Blue condition	Bear condition	Blem condition	Brince condition
Prevoicing 6-Prevoicing 12	t1(59) = .42 not significant	t1(59) = 1.44 not significant	t1(59) = -.351 not significant	t1(59) = -1.22 not significant
	t2(30) = .41 not significant	t2(29) = .99 not significant	t2(33) = -.50 not significant	t2(30) = -.99 not significant
Prevoicing 6-Unrelated	t1(59) = -3.47 p < .001	t1(59) = -5.19 p < .001	t1(59) = -3.27 p < .01	t1(59) = -6.33 p < .001
	t2(30) = -2.97 p < .01	t2(29) = -3.78 p < .001	t2(33) = -2.63 p < .04	t2(30) =-4.86 p < .001
Unrelated-Voiceless	t1(59) = -.14 not significant	t1(59) = -.69 not significant	t1(59) = 1.53 not significant	t1(59) = -1.95 n.s. (p = .056)
	t2(30) = -.50 not significant	t2(29) = -35 not significant	t2(33) = 1.36 not significant	t2(30) = -2.35 p < .04
Prevoicing 0-Prevoicing 6	t1(59) = -.29 not significant	t1(59) = -.45 not significant	t1(59) = 1.07 not significant	t1(59) = 3.98 p < .001
	t2(30) = -.35 not significant	t2(29) = -.35 not significant	t2(33) = .44 not significant	t2(30) = 3.85 p < .001
Prevoicing 0-Voiceless	t1(59) = -4.35 p < .001	t1(59) = -6.52 p < .001	t1(59) = -.99 not significant	t1(59) = -4.31 p < .001
	t2(30) = -5.14 p < .001	t2(29) = -4.44 p < .001	t2(33) = -.89 not significant	t2(30) = -3.8 p < .001

F1(4,236) = 11.50, p < .001, F2(4,120) = 9.82, p < .001 in the Blue condition; and F1(4,236) = 25.76, p < .001, F2(4,116) = 12.15, p < .001 in the Bear condition.

For each condition (and also in each of the Nonword conditions) t-tests on the following five planned comparisons were carried out: prevoicing 12 - prevoicing 6, prevoicing 6 - unrelated, voiceless - unrelated, prevoicing 0 - prevoicing 6, prevoicing 0 - voiceless. We corrected for the number of comparisons in each case by applying a Modified Bonferroni test (Keppel 1982; pp. 148-149), which resulted in a rejection probability of .04 (five comparisons between five conditions were planned). The outcomes of the t-tests for the five planned comparisons within each lexical status condition are shown in Table 4-3. Parallel analyses of arcsine transformed error rates were also performed. The mean error proportions are given in Table 4-4.

The t-tests on the five planned comparisons showed the same pattern in both word target conditions. There was no difference in RTs between targets preceded by primes with 12 or 6 periods of prevoicing. Voiced primes with prevoicing showed a clear priming effect: Lexical decisions in these priming conditions were faster than decisions to targets preceded by unrelated primes. The RTs to targets preceded by voiceless primes were as slow as RTs to targets preceded by unrelated Primes. This confirms that a minimal phonemic change like a change in voicing is sufficient to make the priming effect disappear. A nonword such as *plauw* does not prime responses to the word *blauw* and a word such as *peer* does not prime responses to the word *beer.* Finally, the deletion of prevoicing (prevoicing 0) seems to have no effect on the amount of priming: the words starting with plosives without prevoicing provided the same amount of priming as the same words starting with plosives with prevoicing.

In the analysis of the errors on word targets, there was a main effect of priming (F1(4,236) = 11.12, p < .001; F2(4,236) = 58.41, p < .001), but no effect of competitor and no interaction. The two separate ANOVAs, however, showed that the effect of priming condition in the Blue condition was significant in the participants analysis (F1(4,236) = 3.49, p < .01) but not in the items analysis. In the Bear condition the effect of priming condition was significant both by participants (F1(4,236) = 8.74, p < .001) and by items (F2(4,116) = 8.64, p < .001).

The outcomes of the planned comparisons showed that, in contrast to the RTs, there was a significant effect in proportion of errors between prevoicing 12 and prevoicing 6 (t1(59) = 2.30, p <.04; t2(29) = 2.72, p < .04). Although the proportion of errors is small in both conditions, there were significantly more errors in lexical decisions to word targets preceded by identical primes with 6 periods of prevoicing than to targets preceded by primes with 12 periods of prevoicing. In addition to this, there were significantly more errors in lexical decisions to word targets preceded by voiceless primes than to targets preceded by

Table 4-4. *Percentage of errors in the lexical decision task of Experiment 4.1*

	Blue condition	Bear condition	Blem condition	Brince condition
Prevoicing 12	3.8	2.5	0.7	1.6
Prevoicing 6	3.8	5.6	3.4	3.8
Prevoicing 0	3.2	3.9	2.9	1.9
Voiceless	7.8	12.2	2.2	5.4
Unrelated	3.8	8.1	3.4	4.0

voiced primes without prevoicing ($t(59) = -3.80$, $p < .001$: $t(29) = -4.01$, $p < .001$). However, no difference was found between the prevoicing 6 and prevoicing 0 conditions. In the latter respect the error pattern reflected the RT pattern.

Next, we focus on the conditions with nonword targets (Blem and Brince; see Figure 4-2). There was a different pattern between targets without a competitor and those with a competitor: In addition to a significant priming effect ($F1(4,236) = 19.12$, $p < .001$; $F2(4,252) = 14.66$, $p < .001$) there was a main effect of competitor ($F1(1,59) = 50.56$, $p < .001$; $F2(1,63) = 14.06$, $p < .001$) and a significant interaction between priming condition and competitor ($F1(4,236) = 7.38$, $p < .001$; $F2(4,252) = 6.12$, $p < .001$). The main effect of competitor showed that participants were slower in rejecting a nonword when the nonword had a voiceless word competitor than when the nonword did not have a voiceless word competitor. The two separate ANOVAs showed a significant effect of priming condition in both nonword target conditions: $F1(4,236) = 3.85$, $p < .01$; $F2(4,132) = 2.53$, $p < .05$ in the Blem condition and $F1(4,236) = 20.26$, $p < .001$; $F2(4,120) = 17.25$, $p < .001$ in the Brince condition.

T-tests (see Table 4-3) showed that in the Blem condition the only significant difference was between the prevoicing 6 and unrelated conditions. RTs to targets preceded by primes starting with voiced plosives with prevoicing were faster than RTs to targets preceded by unrelated primes. There was no difference between the RTs to targets preceded by voiceless primes and unrelated primes. This indicates that there was phonological priming from nonword to nonword, but only when the complete string of phonemes of the prime matched the string of graphemes of the target. A nonword such as *blem* primed responses to the nonword *blem*, but a nonword such as *plem* did not prime responses to the nonword *blem*. The effect of phonological priming on nonwords was rather small in comparison to the effect of identity priming on words.

Phonological priming was also present in the Brince condition: the nonword targets were primed by nonword primes starting with prevoicing and there was no difference between prevoicing 6 and prevoicing 12. This time, however, primes starting with voiceless plosives showed an inhibitory effect: listeners were slower in deciding that *brins*, for example, was not a word when they had just heard the word *prins* than when the target was preceded by an unrelated nonword. This difference was significant by items, but not by participants (p = .056). As predicted, the decision that a string of graphemes did not form a word was slowed down somewhat by the activation of a word that was very similar to the nonword.

The most interesting condition is the prevoicing 0 condition. The results show that in the Brince condition there was a difference between voiced primes with and without prevoicing. When a nonword target was preceded by a prime starting with a voiced plosive with prevoicing, participants were faster in rejecting the target as a word than when the target was preceded by a prime starting with a voiced plosive without prevoicing. This suggests that nonwords without prevoicing such as *brins* had activated their matched words (e.g., *prins*). However, RTs to targets preceded by prevoicing 0 primes were still faster than RTs to targets preceded by voiceless primes, indicating, for example, that *brins* without prevoicing had activated the word *prins* to a lesser extent than *prins* itself had.

In the analysis of the errors on nonword targets, there was a main effect of priming condition (F1(4,236) = 3.10, p < .05; F2(4,252) = 3.54, p < .01), but no effect of competitor and no interaction. Nevertheless, in the separate ANOVAs there was no effect of priming condition in the Blem condition, but there was a significant effect in the Brince condition (F1(4,236) = 2.69, p < .05; F2(4,120) = 3.63, p < .01).

The planned t-tests showed that in the Brince condition the only significant difference in error rates was between targets preceded by voiceless primes and targets preceded by voiced primes without prevoicing (t1(59) = -2.35, p < .04; t2(30) = -2.76, p < .01). In line with the RT pattern, participants made more errors in rejecting, for example, *brins* as a nonword when it was preceded by *prins* than when it was preceded by *brins* without prevoicing. There was however no significant difference in errors between targets preceded by the prevoicing 6 and prevoicing 0 primes.

Discussion

We found the following general patterns. First, lexical decisions to word and nonword targets were facilitated by primes which were identical to the targets in comparison to primes which were unrelated. Second, when the first phoneme of the prime only differed from the target in the voicing of the initial plosive, there was no priming effect in comparison to the unrelated prime. Third, the decision that a nonword target was not a word tended to be slowed down in conditions where a voiceless word competitor could have been activated by the prime.

In addition to these general patterns, we observed the following effects of prevoicing variation. There were no differences between lexical decision latencies to targets preceded by primes with 12 or 6 periods of prevoicing in any of the four lexical status conditions. Only in one of the four conditions was there a difference in the proportion of errors between these two conditions. Overall, the difference between 12 and 6 periods of prevoicing before the burst of voiced plosives did not reliably affect the activation of the lexical candidates starting with these plosives. Likewise, the same quantitative difference between six periods of prevoicing and no prevoicing did not affect the lexical decisions to voiced word targets; similar RTs were found for lexical decisions to word targets in the Blue condition preceded by primes with prevoicing and primes without prevoicing.

Interestingly, however, there was a difference in RTs between the prevoicing 6 and prevoicing 0 conditions in one of the lexical status conditions, namely in the Brince condition. When a visual nonword target such as *brins*, which has a voiceless word competitor, was preceded by *brins* without prevoicing, lexical decisions on this target were significantly slower than when it was preceded by *brins* with 6 periods of prevoicing. This suggests that the prime had activated the voiceless word competitor (e.g., *prins*), making it harder to reject the visual nonword target as being a word. In the Blue and Blem conditions there were no voiceless word competitors to be activated, and hence no effect of prevoicing deletion was found in these conditions (i.e., in the absence of a voiceless lexical competitor, there is no interference in making lexical decisions to the voiced target).

In both the Brince and Bear conditions, however, there were voiceless competitors which could have been activated, but the difference between priming conditions with and without prevoicing was only found in the Brince condition. The crucial difference between these two conditions is that in the former condition only the voiceless counterpart was a word, while in the latter condition both the voiced and voiceless counterparts were words. In the Brince condition there was only one possible word candidate (e.g., *prins*, prince), while in the Bear condition there were two word candidates that had to compete with each other (e.g., *beer*, bear, and *peer*, pear). The identification data suggest that items starting with voiced plosives without prevoicing activate both voiced and voiceless word candidates, but that candidates starting with voiced plosives receive more activation than candidates starting with voiceless plosives, since the plosives without prevoicing are more often judged to be voiced than voiceless. This implies that in the Brince condition the unprevoiced primes activated the voiceless word candidates to some extent. This small amount of activation was not suppressed by other lexical candidates, since there were no voiced word candidates, so lexical decisions to the nonword targets were slowed. In the Bear condition, however, both the voiced and voiceless word candidates were activated. Both word candidates would compete for recognition, and based on the identification data, one would expect that the voiced word candidate would suppress the activation of the voiceless candidate and win the competition.

One possible explanation for these findings, therefore, is that the phonological lexical representation of the voiced word candidate was activated so strongly in comparison to the phonological representation of the voiceless word candidate that, at the moment of the lexical decision on the target, the voiced candidate had already suppressed the activation of the voiceless word candidate completely, as if the voiceless candidate had never been activated. An alternative explanation bears on the fact that listeners were asked to make lexical decisions on visual targets. Orthographic lexical representations will receive most of their activation as a result of the visual presentation of the target and only some activation from phonological representations. It is therefore possible that the activation of the phonological voiceless word representation given a Bear prime had not been suppressed completely by the more strongly activated representation of the voiced competitor, but that this activation could not be measured in visual lexical decisions. After all, the visual target unambiguously started with B or D and had activated the orthographic lexical representation of the voiced candidate. Thus, there was clear evidence from the visual modality that the target was voiced. Furthermore, since the prime without prevoicing had activated both voiced and voiceless word candidates, the orthographic lexical representation also received activation from the phonological representation of the voiced lexical candidate. So the orthographic voiced representation has more visual support and more support from the auditory system than the orthographic voiceless representation. Lexical competition in the visual word-recognition system could thus allow the voiced candidate to win, suppressing any activation of the orthographic representation of the voiceless candidate. The voiceless word candidate would then not influence the lexical decision on the voiced target.

If this latter explanation is correct, the use of the voiceless word candidate as the visual target rather than the voiced word candidate might allow the weakly activated phonological word representation of the voiceless word candidate to influence the lexical decision on the visual target. When the target is voiceless there is evidence from the visual modality that the word target is voiceless, and the weakly activated phonological representation of the voiceless lexical candidate might now be able to add activation to its orthographic representation. We should then find faster lexical decisions to voiceless targets when they are preceded by voiced primes without prevoicing than when they are preceded by primes with prevoicing. If the first explanation is correct, however, and the activation of the voiceless word candidate in the auditory system is already completely suppressed by the more strongly activated voiced word candidate, then there should be no difference in lexical decisions to voiceless word targets preceded by voiced word primes without prevoicing and voiced word primes with prevoicing. Experiment 4.2 was designed to test these two alternative explanations.

EXPERIMENT 4.2

In this experiment participants heard the same primes as in Experiment 4.1, but this time the targets on which they had to perform the lexical decision task were not the voiced members of the selected pairs, but the voiceless members (see Table 4-1). On the basis of the results of Experiment 4.1 we could make detailed predictions for the different combinations of lexical status conditions and priming conditions in Experiment 4.2. Moreover, the design of Experiment 4.2 gave us the opportunity to gain more insight in the effects of prevoicing variation which were found in Experiment 4.1, especially in the competition process between the voiced and voiceless word candidates in the Bear condition after the presentation of voiced primes without prevoicing.

We expected to find the same general patterns as were found in Experiment 4.1. However, since in Experiment 4.2 the voiceless members of the pairs were presented as targets, the patterns were now expected to appear in different conditions than in Experiment 4.1. First, the identity priming effect for both word and nonword targets was now expected to appear when the voiceless targets were preceded by voiceless primes. Second, when targets were preceded by voiced primes (meaning that primes and targets differed in the voicing of the initial plosive), no priming effect was expected. Third, when the target was a nonword and the prime was a word which only differed from the target in initial voicing, an inhibitory effect was expected. In Experiment 4.2, this would be the case in the Blue condition, when a voiceless nonword target such as *plauw* was preceded by a voiced word prime such as *blauw.*

In addition to these general patterns, we again expected to find the following effects of prevoicing variation. As in Experiment 4.1, we expected to find no difference between the prevoicing 12 and prevoicing 6 conditions. We have argued that primes with 12 periods of prevoicing activate voiced word candidates (and voiceless word candidates) to the same degree as those with 6 periods of prevoicing. We therefore did not expect to find a difference between the effects of these primes in lexical decision latencies to voiceless targets in the conditions where the voiceless members were words (the Bear and Brince conditions). Furthermore, we did not expect to find differences between either the prevoicing 12 and prevoicing 6 conditions or the prevoicing 6 and prevoicing 0 conditions in the lexical status conditions where the voiceless members were nonwords (i.e., in the Blem and Blue conditions). In the Blem condition there should be no competition at all since both members were nonwords; in the Blue condition the voiced words should be activated, but there are no voiceless word candidates to be activated. So in both of these conditions there should be no competition and no effect of prevoicing variation.

However, when the voiceless members did form words (i.e., in the Brince and Bear conditions) we expected to see a difference between primes with prevoicing and primes without prevoicing, since the latter primes should activate the voiceless word candidates more strongly. In the Brince condition the voiceless word candidate was the only candidate and therefore the primes without prevoicing were expected to facilitate lexical decisions on the voiceless targets.

The crucial condition was the Bear condition, in which both voiced and voiceless counterparts were words. Primes without prevoicing should activate the voiceless word candidates, as was indicated by the results in the Brince condition in Experiment 4.1. The results of the Bear condition in Experiment 4.1, however, did not show any indication that this was indeed the case. We have already proposed two possible explanations for the failure to find a difference in this condition between the prevoicing 6 and prevoicing 0 conditions. If it is the case that the more strongly activated voiced candidate had immediately suppressed the activation of the voiceless candidate completely (i.e., as if it never had been activated) we should expect to again find no difference between the prevoicing 6 and prevoicing 0 priming conditions with voiceless targets. In contrast, if it is the case that the activation of the voiceless word competitor was not completely suppressed when a prevoicing 0 stimulus was heard, but that the lack of difference between these two conditions was due to lexical competition in the visual word-recognition system, we should expect that the activation of the phonological word representation of the voiceless candidate would be able to influence lexical decisions to the voiceless word targets. That is, we would then expect that the lexical decisions on voiceless word targets in the Bear condition would be faster when targets were preceded by primes without prevoicing than when they were preceded by primes with prevoicing.

Method

Participants

Sixty volunteers from the Max Planck Institute participant pool took part in this experiment, in exchange for a small payment. All were native speakers of Dutch and none reported any hearing loss. None had participated in Experiment 4.1.

Materials and Procedure

The materials were identical to those used in Experiment 4.1, except for the visual targets. This time, the voiceless member rather than the voiced member of each pair served as target. The target was now identical to the prime in the voiceless prime condition and differed in the voicing of the initial plosive from the primes in the three prevoicing conditions. Note that, in comparison to Experiment 4.1, the use of the voiceless counterparts as targets resulted in a change in the lexical status of the targets in the Brince and Blue conditions. As a result, the lexical status of the unrelated primes was not identical to the lexical status of the related voiced primes in these conditions. Table 4-1 shows the complete design of Experiment 4.2 with an example for each combination of priming condition and lexical status condition. The same filler pairs were presented as in Experiment 4.1. The instructions and procedures were identical to those for the lexical decision task in the previous experiment. There was, however, no phoneme identification task.

Results

The mean RTs of the correct responses to targets in each priming condition are plotted separately for each of the four lexical status conditions in Figure 4-3 and Figure 4-4. Note that this time the division based on correct “yes” and “no” responses resulted in different combinations of lexical status conditions than in Experiment 4.1. Since the lexical status of the target is now determined by the lexical status of the voiceless counterpart, “yes” responses were correct in the Brince and Bear conditions (Figure 4-3) and “no” responses were correct in the Blem and Blue conditions (Figure 4-4). The analyses of the RTs and the errors

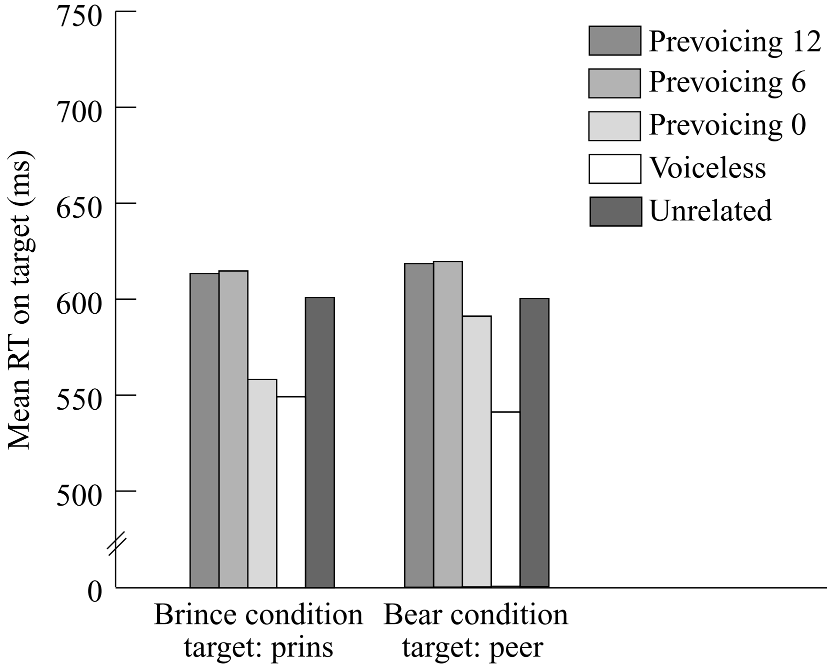

Figure 4-3. *Mean RTs to voiceless word targets in the lexical decision task of Experiment 4.2*

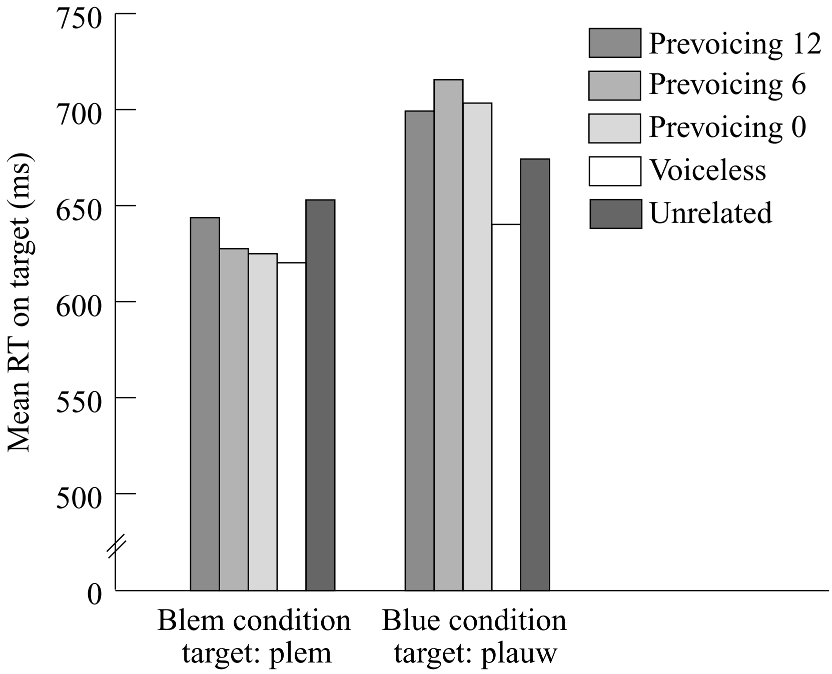

Figure 4-4. *Mean RTs to voiceless nonword targets in the lexical decision task of Experiment 4.2*

Table 4-5. *T-tests for planned comparisons in the lexical decision task for each of the five comparisons in each of the four lexical status conditions in Experiment 4.2*

Comparison	Brince condition	Bear condition	Blem condition	Blue condition
Prevoicing 12-Prevoicing 6	t1(59) = -.17 not significant	t1(59) = -.00 not significant	t1(59) = -1.43 not significant	t1(59) = 1.05 not significant
	t2(30) = .11 not significant	t2(29) = .28 not significant	t2(33) = -1.35 not significant	t2(30) = 1.34 not significant
Prevoicing 6-Unrelated	t1(59) = .85 not significant	t1(59) = 1.70 not significant	t1(59) = -2.49 p < .04	t1(59) = 3.03 p < .01
	t2(30) = .81 not significant	t2(29) = 1.88 not significant	t2(33) = -1.63 not significant	t2(30) = 2.57 p < .04
Unrelated-Voiceless	t1(59) = 4.57 p < .001	t1(59) = 6.99 p < .001	t1(59) = 4.51 p < .001	t1(59) = 3.12 p < .01
	t2(30) = 4.52 p < .001	t2(29) = 4.70 p < .001	t2(33) = 2.20 p < .04	t2(30) = -.3.15 p < .01
Prevoicing 0-Prevoicing 6	t1(59) = -6.23 p < .001	t1(59) = -2.36 p < .04	t1(59) = -.71 not significant	t1(59) = -1.06 not significant
	t2(30) = -5.35 p < .001	t2(29) = -2.32 p < .04	t2(33) = -.28 not significant	t2(30) = -.85 not significant
Prevoicing 0-Voiceless	t1(59) = .67 not significant	t1(59) = 5.20 p < .001	t1(59) = .53 not significant	t1(59) = 6.73 p < .001
	t2(30) = .84 not significant	t2(29) = 3.71 p < .001	t2(33) = .39 not significant	t2(30) = -3.54 p < .001

were identical to those carried out in Experiment 4.1. The outcomes of all separate t-tests on the RTs are shown in Table 4-5 and the mean percentage of errors are shown in Table 4-6.

As before, we will first discuss the RTs and the errors of the responses to word targets. The ANOVAs on both the Brince and Bear conditions showed that there was a significant effect of priming condition: F1(4,236) = 18.21, $p < .001$; F2(4,236) = 20.55, $p < .001$, but no significant effect of competitor (whether the voiced counterpart of the voiceless target was a word or not). The interaction between priming condition and competitor was significant in the participants analysis (F1(4,236) = 17.67, $p < .001$), but not in the items analysis. Individual ANOVAs showed that the effect of priming condition was significant in both lexical status conditions: F1(4,236) = 16.97, $p < .001$, F2(4,120) = 11.05, $p < .001$ in the Brince condition; and F1(4,236) = 16.67, $p < .001$, F2(4,116) = 10.75, $p < .001$ in the Bear condition.

Consider first the RTs of the responses to word targets in the Brince condition. As in all lexical status conditions in Experiment 4.1, there was no difference between RTs to targets preceded by primes with 6 periods of prevoicing and those preceded by primes with 12 periods of prevoicing. The RTs in those two voiced priming conditions were not different from RTs in the unrelated conditions. This indicates that responses to the word target *prins,* for example, were not facilitated by the presentation of the nonword prime *brins*. As in Experiment 4.1, where responses to a word target such as *blauw* were not facilitated by the nonword prime *plauw,* words were not primed by primes which were identical to the targets except for the voicing of the initial plosive. The voiceless prime showed a clear priming effect: participants were faster in deciding that the target was a word when the prime was identical to the target. There was no difference in RTs between responses to targets preceded by primes without prevoicing (prevoicing 0) and voiceless primes. This indicates that nonwords such as *brins* without prevoicing had activated their matched words (e.g., *prins*). Responses to targets preceded by these primes were faster than responses to targets preceded by voiced primes.

As shown in Figure 4-3 and Table 4-5, the pattern in the Bear condition is similar to the pattern in the Brince condition, except for the effect of voiced primes without prevoicing. There was again no difference between the prevoicing 12 and prevoicing 6 conditions, and there was no difference in RTs between responses to the voiceless word targets preceded by these voiced primes and unrelated primes. Again, the voiceless prime showed a clear priming effect: RTs were faster to targets preceded by voiceless primes than by unrelated primes. In contrast to the Brince condition, in which primes without prevoicing provided the same amount of priming as the voiceless primes, responses to targets preceded by primes without prevoicing were slower than responses to targets pre-

Table 4-6. *Percentage of errors in the lexical decision task of Experiment 4.2*

	Brince condition	Bear condition	Blue condition	Blem condition
Prevoicing 12	8.4	12.0	5.7	2.5
Prevoicing 6	7.8	13.6	4.9	2.5
Prevoicing 0	3.2	5.8	4.5	3.7
Voiceless	1.6	3.3	4.3	3.2
Unrelated	5.1	4.7	5.1	3.7

ceded by voiceless primes, but faster than responses to targets preceded by voiced primes. Prevoicing 0 was thus an intermediate case between unambiguously voiced and unambiguously voiceless primes.

The analysis of the errors on word targets showed that there was a main effect of priming condition (F1(4,236) = 17.33, p < .001; F2(4,236) = 16.68, p < .001). There was no effect of competitor and the interaction between priming condition and competitor was only significant in the participants analysis (F1(4,236) = 4.50, p < .01). The two individual ANOVAs showed that in both the Brince and the Bear conditions there was a significant effect of priming condition (F1(4,236) = 7.88, p < .001; F2(4,120) = 8.74, p < .001 in the Brince condition; and F1(4,236) = 11.07, p < .001; F2(4,116) = 8.80, p < .001 in the Bear condition).

The t-tests showed that in the Brince condition the proportion of errors was lower when the voiceless targets were preceded by voiceless primes than when preceded by unrelated primes (t1(59) = 2.80, p < .01; t2(30) = 3.52, p < .001). The priming effect in RTs of primes which were identical to the targets was therefore also reflected in the error pattern. In addition to this, there were fewer errors when the target was preceded by a voiced prime without prevoicing than when it was preceded by a voiced prime with prevoicing (t1(59) = -3.38, p < .001; t2(30) = -4.47, p < .001). This finding is also in line with the RT pattern.

In the Bear condition, there were significantly more errors when the voiceless word target was preceded by a voiced word prime than when this target was preceded by an unrelated prime (t1(59) = 4.04, p < .001; t2(29) = 3.20, p < .01). This reflects the difference in RTs which was found between prevoicing 6 and unrelated primes. In contrast to the RT pattern, which showed that prevoicing 0 was an intermediate case between voiced and voiceless primes, the error pattern showed that there were fewer errors in the prevoicing 0 than in the prevoicing 6 condition (t1(59) = -3.51, p < .001; t2(29) = -3.00, p < .01), but no difference between the prevoicing 0 and voiceless conditions.

Next, we discuss the RTs and errors to nonword targets (i.e., the Blem and Blue conditions; see Figure 4-4 and Table 4-4). The combined ANOVAs on these conditions showed a significant main effect of priming condition (F1(4,236) = 31.38, $p < .001$; F2(252) = 5.92, $p < .001$), a significant main effect of competitor (F1(1,59) = 6.49, $p < .05$; F2(1,63) = 19.63, $p < .001$), and a significant interaction between priming condition and competitor (F1(4,236) = 16.22, $p < .001$; F2(4,252) = 4.82, $p < .001$). The two separate ANOVAs showed that in the Blem condition the effect of priming condition was significant in the participants analysis (F1(4,236) = 4.87, $p < .001$), but not in the items analysis. In the Blue condition the effect was significant in both the participants and items analyses (F1(4,236) = 11.69, $p < .001$; F2(4,120) = 7.93, $p < .001$).

The outcomes of the t-tests showed that in the Blem condition only one of the planned comparisons was significant. Responses to targets preceded by voiceless primes were significantly faster than responses to targets preceded by unrelated primes, indicating that there was identity priming. There was no significant difference between the prevoicing 0 and voiceless conditions.

In the Blue condition there was again no difference between the prevoicing 6 and prevoicing 12 conditions. Furthermore, responses to voiceless nonword targets preceded by voiced word targets were slower than responses to these targets preceded by unrelated primes. This indicates that lexical decisions to nonword targets such as *plauw* were inhibited by word primes such as *blauw* in comparison to unrelated word primes. As before, there was a clear priming effect of the voiceless primes: RTs were faster when targets were preceded by voiceless primes than when preceded by unrelated primes. The voiced primes without prevoicing behaved like the other voiced primes: there was no difference between prevoicing 0 and prevoicing 6, and RTs to targets preceded by prevoicing 0 primes were significantly slower than RTs to targets preceded by voiceless primes.

The error analyses on nonword targets showed that there were no significant main effects and no significant interactions. The proportion of errors was not affected by the type of prime in either of the nonword target conditions.

Discussion

Experiment 4.2 showed the same general patterns as Experiment 4.1. First, all priming conditions in which primes were identical to targets showed facilitation in the lexical decisions on the targets. In Experiment 4.1 this effect appeared in all voiced priming conditions (prevoicing 6 and prevoicing 12) and in Experiment 4.2 in all voiceless priming conditions in comparison to the unrelated priming conditions. In other words, when listeners saw a word or nonword on the screen and had to decide whether that target was a word or a nonword, they found it easier to make a decision when they had heard the auditory version of the written target than when they had heard an unrelated prime. Second, the identity priming effect only appeared when the prime was identical to the target. If the prime differed from the target only in the voicing of the initial plosive, the facilitation effect completely disappeared. This effect went both ways: responses to targets starting with voiced plosives were not facilitated by the voiceless counterparts of the targets (the voiceless priming conditions in Experiment 4.1) and responses to targets starting with voiceless plosives were not facilitated by the voiced counterparts of the targets (the voiced priming conditions in Experiment 4.2). Third, making a nonword decision was harder when the visual nonword target was preceded by an auditory counterpart with the opposite voicing, but only when the prime was a word (the voiceless priming condition in the Brince condition of Experiment 4.1 and the voiced priming conditions in the Blue condition of Experiment 4.2).

As in Experiment 4.1, there were no differences in RTs between the priming conditions with 6 periods of prevoicing and the priming conditions with 12 periods of prevoicing. This indicates that both these types of prevoicing variation provide the same degree of evidence that the plosive is voiced. As a result, primes starting with plosives with 12 or 6 periods of prevoicing activate lexical candidates equally strongly. In addition, there was again no difference between primes with and without prevoicing when the voiceless members of the pairs were nonwords. This indicates that voiced primes without prevoicing contain sufficient acoustic support for the primes to be considered to begin with voiced plosives. In the conditions where the voiceless members were words, however, the presentation of primes without prevoicing resulted in different RTs in comparison to primes with prevoicing. The following differences were observed: slower lexical decisions to voiced nonword targets after hearing prevoicing 0 primes than after prevoicing 6 primes (the Brince condition of Experiment 4.1); faster lexical decisions to voiceless word targets after prevoicing 0 primes than after prevoicing 6 primes (the Brince condition of Experiment 4.2); and faster lexical decisions to voiceless word targets after prevoicing 0 primes than after prevoicing 6 primes (the Bear condition of Experiment 4.2). These differences indicate that, in addition to voiced word candidates, voiceless word candidates are also activated by primes without prevoicing.

The only condition in which the activation of the voiceless word candidate was not detectable was in the Bear condition of Experiment 4.1. We proposed two possible explanations for the absence of any visible competition between the two candidate words in this condition. The first was that the voiced candidate had immediately suppressed the activation of the voiceless candidate, such that at the moment of the lexical decision the latter was completely suppressed. The second was that the voiceless candidate was still active, but that the combined evidence in favor of the voiced competitor from the auditory and visual modalities was so strong that the weakly activated voiceless lexical candidate in the auditory modality could not affect visual lexical decision latencies. Critically, in the Bear condition of Experiment 4.2, in which the voiceless words served as targets, there was a significant difference between RTs to targets preceded by primes with prevoicing and primes without prevoicing. Participants were faster in deciding that a word such as *peer* was a word when it was preceded by *beer* without prevoicing than when it was preceded by *beer* with prevoicing. This result suggests that the voiceless word candidate was still activated at the moment of the lexical decision, but that it depends on the voicing of the initial grapheme of the visual target whether this activation can influence lexical decision latencies. When the prime was voiced but without prevoicing and the target was also voiced, the small activation of the voiceless word candidate in the auditory system had no chance to influence responses to the visual target. However, when the prime was voiced without prevoicing and the target was voiceless, the small activation of the voiceless word candidate could speed up lexical decisions on the voiceless target. This will be discussed in more detail in the General Discussion.

In summary, Experiments 4.1 and 4.2 show that variation in prevoicing affects the activation of phonological representations of words. There was no effect of the duration of prevoicing when there was prevoicing. Both tokens with 6 periods of prevoicing and tokens with 12 periods of prevoicing were clear voiced plosives and items starting with these plosives activated word candidates starting with voiced plosives equally strongly. Primes without prevoicing, however, showed a different pattern than primes with prevoicing. Removing the prevoicing of voiced plosives took away an important cue for the voicing of that plosive, but the remaining cues provided sufficient evidence for the plosive to be more voiced than voiceless. Nevertheless, items starting with these plosives without prevoicing activated word candidates starting with voiced plosives and candidates starting with voiceless plosives. In line with the fact that the plosives without prevoicing were more voiced than voiceless, word candidates starting with voiced plosives received more activation than word candidates starting with voiceless plosives.

The results so far show that the same quantitative difference in prevoicing of 12 versus 6 periods of prevoicing and 6 versus 0 periods of prevoicing has qualitatively very different effects on lexical access. This qualitative difference is in line with the fact that, although all three prevoicing durations fall within the natural range of prevoicing variation, plosives with prevoicing are unambiguously voiced while plosives without prevoicing are not. The absence of a difference in the identification and priming tasks between the two conditions with prevoicing, however, could also been due to a lack of perceptual sensitivity, that is, listeners may be unable to hear the difference between 12 and 6 periods of prevoicing. The possibility remains that, although the difference between 6 and 0 periods of prevoicing is detectable, the same quantitative difference between 12 and 6 periods of prevoicing is not. Before we can conclude that the difference between 12 and 6 periods of prevoicing does not affect lexical access, we first have to establish that listeners can actually hear the difference between these two durations of prevoicing.

Experiment 4.3 was designed to test the ability of listeners to discriminate between the three different durations of prevoicing. In addition to this, our aim was to show that the effects of prevoicing variation on lexical access which were found in the two priming experiments would also appear when using a different experimental paradigm, namely, discrimination.

EXPERIMENT 4.3

In this experiment listeners were asked to indicate whether two auditory items were the same or different. The voiced primes of the previous experiments were used in this discrimination task. The two items were either identical, or differed only in the duration of prevoicing of the initial voiced plosive. Given the effects of the two priming studies, we predicted that listeners would be able to discriminate between items with and without prevoicing (prevoicing 12 - prevoicing 0 and prevoicing 6 - prevoicing 0). The crucial question was whether listeners could also discriminate between items with two different durations of prevoicing (prevoicing 12 - prevoicing 6). One would expect that it is easier to detect the difference between the absence and presence of a particular acoustic event (prevoicing) than to detect the difference between the same acoustic events which only differ in duration. We therefore predicted that it would be harder to discriminate between prevoicing 12 and prevoicing 6 than between prevoicing 12 and prevoicing 0 or between prevoicing 6 and prevoicing 0.

Furthermore, given the results so far, we also expected that the different activation patterns of lexical candidates across conditions would play a role in the discrimination task. Although listeners were instructed to focus on small acoustic differences between the two stimuli, we assumed that they would make use of any available information that might help them to perform the task. We therefore expected that discriminating between two items would be facilitated by any support from the lexical level. On the basis of the results of Experiments 4.1 and 4.2, it appears that voiced plosives without prevoicing activate the prelexical representations corresponding to both voiced and voiceless phonemes, and thus that items starting with these plosives activate lexical candidates with initial voiced plosives and lexical candidates with initial voiceless plosives. If this is indeed the case, we should expect that these prelexical and lexical activations would both influence discrimination performance. Items without prevoicing would activate both voiced and voiceless prelexical representations, while items with prevoicing would preferentially activate voiced representations. Similarly, if the items had voiceless word competitors, items without prevoicing should activate the voiceless word candidates to a significant extent (and voiced candidates if there were any), while items with prevoicing should not activate voiceless word candidates to any significant extent. We therefore predicted that the presence of a voiceless word candidate would facilitate the discrimination between items with and without prevoicing.

Method

Participants

Twelve volunteers from the Max Planck Institute participant pool were paid for taking part. They were all native speakers of Dutch. None reported any hearing loss and none had participated in any of the previous experiments.

Materials

The materials consisted of the 160 voiced primes of Experiments 4.1 and 4.2. For each item, each of the three prevoicing conditions were used (12, 6 and 0 periods of prevoicing). For each item six pairs were constructed in such a way that all combinations of prevoicing (prev) appeared: prev 12 - prev 12, prev 6 - prev 6, and prev 0 - prev 0 (the “same” pairs); and prev 12 - prev 0, prev 6 - prev 0, and prev 12 - prev 6 (the “different” pairs). The order of prevoicing conditions within the different pairs were balanced, such that the members of half of the pairs were presented in one order (e.g., longer prevoicing durations first), and the members of the other half of the pairs were presented in the reversed order (e.g., longer prevoicing durations second). In total there were 960 pairs.

Procedure

The two members of each pair were separated by an inter-stimulus interval of 300 ms. The interval between the offset of a pair and the onset of the following pair was 1500 ms. Participants were asked to listen carefully, especially to the very beginning of each stimulus, and then indicate, by pressing one of two appropriately labeled buttons, whether the two members sounded exactly the same or were different. Before the experiment started listeners heard 12 “different” pairs and were told beforehand that all these pairs consisted of members which had slightly different onsets. The familiarization phase was followed by a practice phase consisting of 24 pairs. Participants then heard all 960 pairs in random order. There were four different randomized versions, each heard by three listeners.

Results

First, following Macmillan and Creelman (1991), mean d's for each type of “different” pair were calculated for each subject to find out whether listeners could hear the difference between the three different durations of prevoicing. Figure 4-5 shows the mean d's for the three prevoicing combinations. There is clearly a difference between the three combinations. This was confirmed by a one-way ANOVA: F1(2,22) = 32.43, $p < .001$. Tukey HSD tests showed that all pairwise differences between pairs were significant. Subjects found it easiest to discriminate prevoicing 0 and prevoicing 12, whereas discrimination of prevoicing 0 and prevoicing 6 was more difficult, and the difference between prevoicing 6 and prevoicing 12 was the hardest to detect.

All d's differed significantly from zero (prev 12 - prev 0: t(11) = 12.23, $p < .001$; prev 6 - prev 0: t(11) = 11.82, $p < .001$; prev 12 - prev 6: t(11) = 5.92, $p < .001$). This indicates that listeners performed above chance. Moderate performance implies that d' is near unity (Macmillan & Creelman, 1991). The d's of the

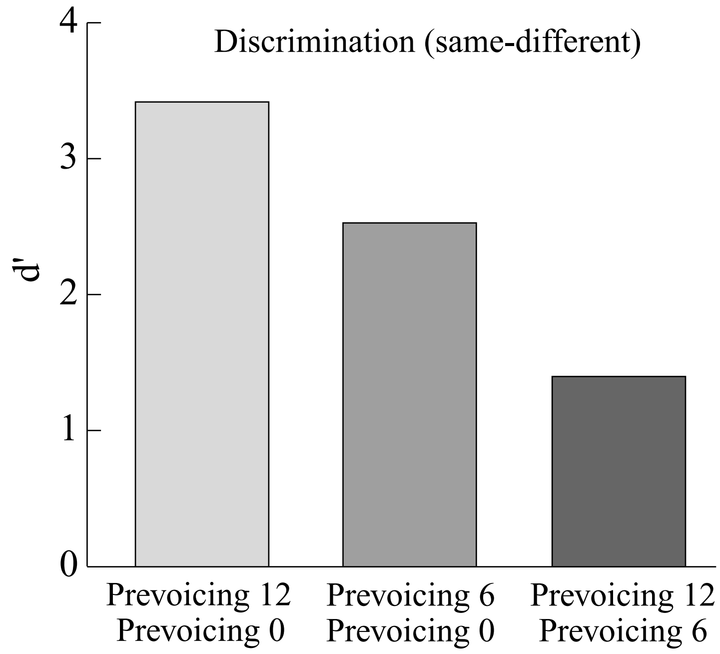

Figure 4-5. *Mean d's for the three combinations of prevoicing in the discrimination task of Experiment 4.3*

pairs prev 12 - prev 0 and prev 6 - prev 0 also differed significantly from unity (prev 12 - prev 0: t(11) = 8.66, p < .001; prev 6 - prev 0: t(11) = 7.15, p < .001). These results thus show that listeners could discriminate between members of all three types of pair, and that discrimination performance was ranked as shown in Figure 4-5.

The influence of the lexical status of both the voiced item and the voiceless counterpart (whether there was a voiceless word competitor or not) was then explored. Therefore, the proportion of correct responses was calculated for each listener as a function of prevoicing pair (prevo 12 - prev 12, prev 6 - prev 6, prev 0 - prev 0, prev 12 - prev 0, prev 6 - prev 0 and prev 12 - prev 6) and lexical status condition (Blue, Bear, Brince and Blem). The mean proportion of correct responses, pooled across subjects, are shown in Figure 4-6 (“same” pairs) and Figure 4-7 (“different” pairs). Note that d' analyses were inappropriate for these more detailed comparisons, since d's are derived from the performance on both the “same” and the “different” pairs. The assumption in the computation of d' is made that the certainty of hearing two stimuli as the same is equal for all “same” pairs (Macmillan & Creelman, 1991). But, as shown below, the percentage of “same” responses differed among the three “same” pairs.

The mean proportions of correct responses for each combination of lexical status and prevoicing pair were converted through an arcsine transformation (Studebaker, 1985) for each participant. Then, both subjects and items three-way repeated measures ANOVAs were performed on the transformed proportions, with prevoicing pair, lexical status of the voiced items and lexical status of the voiceless counterparts as factors. The “same” and “different” pairs were analyzed separately.

In the “same” pairs there was only a main effect of prevoicing pair (F1(2,22) = 7.87, p < .01; F2(2,244) = 76.29, p < .001). A post hoc Tukey HSD test showed that the proportion of correct responses was smaller for prev 0 - prev 0 than for prev 6 - prev 6 or prev 12 - prev 12 in both the subjects and items analysis. There was no difference between the pairs with prevoicing. There were no effects of lexical status of either the voiced items or the voiceless counterparts.

In contrast, in the “different” pairs, there was a main effect of both prevoicing pair (F1(2,22) = 40.43, p < .001; F2(2,244) = 309.67, p < .001) and lexical status of the voiceless counterpart (F1(1,11) = 29.72, p < .001; F2(1,122) = 21.51, p < .001). In addition to these two main effects, there was a significant interaction between those two factors: F1(2,22) = 7.91, p < .01; F2(2,244) = 4.57, p < .05. No other main effects or interactions were significant. A post hoc Tukey HSD test showed that all overall pairwise differences between the three

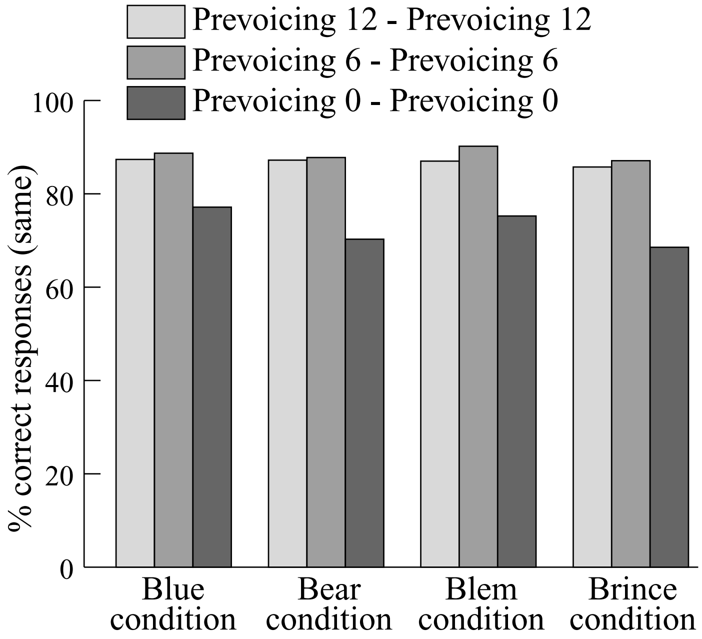

Figure 4-6. *Mean percentage of correct responses ("same") in Experiment 4.3 in each of the pairs of which the two members were identical, plotted separately for each of the four lexical status conditions*

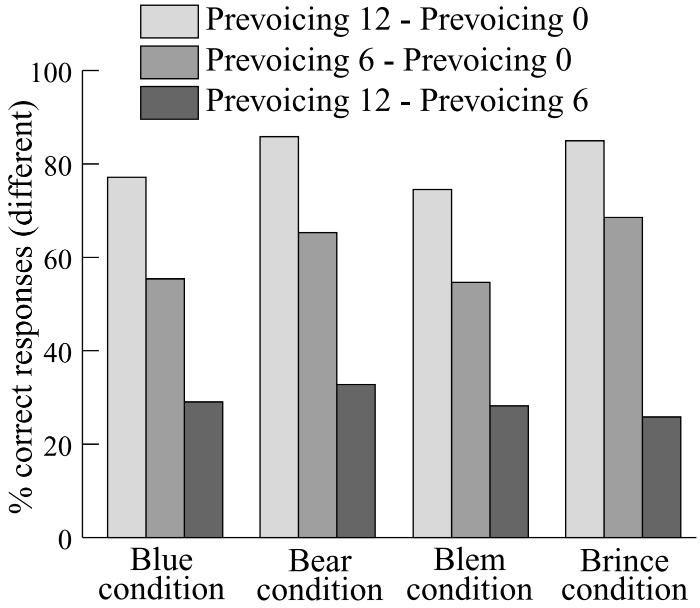

Figure 4-7. *Mean percentage of correct responses ("different") in Experiment 4.3 in each of the pairs of which the two members differed in the prevoicing duration, plotted separately for each of the four lexical status conditions*

pairs were significant at the .05 level (i.e., % correct performance was ranked as follows: prev 12 - prev 0 > prev 6 - prev 0 > prev 12 - prev 6). This pattern confirms what was observed in the d' analysis.

A Tukey HSD test on the six means of the interaction between prevoicing pair and lexical status of the voiceless counterpart showed that there was only a significant difference between word and nonword voiceless counterparts for the pairs prev 12 - prev 0 and prev 6 - prev 0. That is, performance was significantly better in the two pairs which contained one item without prevoicing (the light gray bars and dark grey bars in Figure 4-7) when the voiceless counterpart was a word (Bear and Brince) than when it was a nonword (Blue and Blem). For the prev 12 - prev 6 pairs (the black bars in Figure 4-7), on the other hand, there was no difference in the percentage correct responses as a function of lexical status of the voiceless counterpart. In other words, it was only in the “different” pairs in which one of the members did not have prevoicing that the lexical status of the voiceless counterpart influenced the proportion of correct responses. For those pairs, listeners responded correctly more often when the voiceless counterpart was a word.

Discussion

The results show that listeners were able to detect the differences among all three prevoicing conditions. As predicted, the difference between two items was easier to detect when one of the two items had no prevoicing than when both items had prevoicing. For the pairs with one unprevoiced item, discrimination was easier when the other item had 12 periods of prevoicing than when the other item had 6 periods of prevoicing. The most difficult difference to detect was between 12 and 6 periods of prevoicing. In more than 70% of the cases, listeners indicated that items with these durations of prevoicing were the same. Nevertheless, the d's for these pairs were significantly different from zero, which indicates that listeners performed better than chance.

As mentioned earlier, the finding that it was easier to detect the difference between no prevoicing and prevoicing than between two different durations of prevoicing could be explained purely auditorily. It is easier to detect the difference between the presence and absence of an acoustic event than between different durations of the same event. The outcomes of the discrimination performance on the “same” pairs, however, suggest that there is also another explanation for this finding. Recall that more mistakes were made in labeling the prevoicing 0 “same” pairs than in labeling the prevoicing 6 “same” pairs or the prevoicing 12 “same” pairs. Why would the sameness of the items without prevoicing be harder to detect auditorily than the sameness of the items with prevoicing? The difference between those "same" pairs can be explained by the activation patterns of the prelexical representations: Plosives without prevoicing activate both voiced and voiceless plosive representations, while plosives with prevoicing strongly activate only voiced plosive representations. Thus, two prevoiced plosives are more likely to sound the same than two unprevoiced plosives. These different activation patterns could have facilitated the discrimination within "different" pairs consisting of one item with prevoicing and one without prevoicing. While the activation of both voiced and voiceless plosives upon hearing an unprevoiced plosive hinders the decision that two unprevoiced items are the same, it facilitates the decision that an unprevoiced item and a prevoiced item are different.

Furthermore, the results show a clear effect of the lexical status of the voiceless counterpart. In the conditions with voiceless word competitors (Bear and Brince conditions) fewer mistakes were made in the "different" pairs when one of the pairs had no prevoicing (prev 12 - prev 0 and prev 6 - prev 0) than in the conditions without voiceless word competitors (Blue and Blem conditions). There was no difference between the conditions with and without voiceless word competitor for the "different" pairs with prevoicing (prev 12 - prev 6). The explanation for this is that, upon hearing an item with prevoicing, the voiceless word candidate was not strongly activated, while upon hearing the same item without prevoicing the voiceless candidate was more substantially activated. In other words, in addition to the support from the auditory level and the prelexical level that items with and without prevoicing are different, the activation of the voiceless word candidate provides extra support that the two items are different. This made it easier to discriminate between items with and without prevoicing when there was a voiceless word competitor. For example, it was easier for listeners to hear the difference between *beer* with 6 periods of prevoicing and *beer* without prevoicing (because of the activation of *peer* in the latter case) than to hear the difference between *blauw* with 6 periods of prevoicing and *blauw* without prevoicing (because *plauw* is not a word). This explanation is in line with our earlier findings that primes starting without prevoicing activate voiceless word candidates. The finding that there was no effect of the lexical status of the voiced counterparts on discrimination performance also corresponds well with the results of the priming experiment. The priming results showed that prevoicing variation did not strongly affect the degree of activation of the voiced counterparts. The voiced representations were strongly activated by all three prevoicing variations, including when prevoicing was absent. For this reason, the lexical status of the voiced counterpart did not help listeners discriminate among the three variations of prevoicing.

GENERAL DISCUSSION

The purpose of the present study was to investigate the effect on lexical access of prevoicing variation in Dutch initial plosives. The results of the cross-modal identity priming experiments showed no latency differences between the effects of primes starting with plosives with 12 periods of prevoicing and those of items starting with plosives with 6 periods of prevoicing. In both cases the primes resulted in facilitation of lexical decisions to identical voiced targets in comparison to unrelated targets (Experiment 4.1). Apparently, these degrees of prevoicing result in clearly voiced plosives, and primes starting with either of these plosives activate identical targets equally strongly. The results of the identification phase in Experiment 4.1 indeed showed that these plosives are unambiguously voiced. Items starting with these plosives therefore do not substantially activate word candidates with initial voiceless plosives, as was shown in Experiment 4.2. There was no difference in lexical decisions to voiceless targets preceded by primes with 12 periods and the same targets preceded by primes with 6 periods of prevoicing; in both cases there was no facilitation in comparison to the unrelated priming condition. This pattern of results may have been the result of an inability of listeners to hear the difference between 12 and 6 periods of prevoicing. The discrimination experiment indicated, however, that listeners were in fact capable of distinguishing these two prevoicing durations from each other, so it was not the case that this type of variation was simply not detectable.

Interestingly, the same quantitative difference of six periods of prevoicing did influence lexical access when the prevoicing varied from 6 periods to 0 periods of prevoicing. Deleting the prevoicing takes away an important cue that the plosive is voiced. Although the identification results showed that in general the plosives without prevoicing were still perceived as voiced, the percentage of voiced responses was lower than for the plosives with prevoicing. Nevertheless, the results of Experiments 4.1 and 4.2 indicate that primes starting with voiced plosives without prevoicing activated both lexical candidates starting with voiced plosives and lexical candidates starting with voiceless plosives. This could only happen if, at the prelexical level, both the phonological representation of the voiced plosive and the phonological representation of the voiceless plosive were activated. The acoustic information in these primes without prevoicing was thus somewhat ambiguous; although the remaining acoustic cues favored a voiced plosive, the absence of prevoicing also resulted in activation of the voiceless plosive.

The resulting RT pattern of lexical decisions to targets preceded by these ambiguous primes varied as a function of lexical status condition and target type (voiced or voiceless). In order to understand these different patterns it is necessary to consider the underlying components of the cross-modal identity priming task. In this task, participants are asked to decide as quickly as possible whether the visual target is a word or not. The decision is therefore based on the degree of lexical activation in the visual word-recognition system. The degree of match between the visual input and orthographic lexical representations largely determines the activation pattern in the visual system. In addition, however, the phonological representations which are activated as a result of the auditory input will also spread activation to the visual system. Therefore, participants can make faster word decisions when the visual word target is preceded by an auditory prime which is identical to the target than when it is preceded by a phonologically unrelated prime. If a nonword target is preceded by a word prime that is very similar to the target, lexical activation in the auditory system will spread activation to the visual system, which will result in slower nonword decisions. Therefore, both positive and negative lexical decision latencies reflect the degree of lexical activation in both the visual and the auditory systems if prime and target are phonologically similar. Furthermore, in both systems there will be competition between activated lexical candidates. Since phonological representations that are activated as a result of the acoustic input will in turn activate matching orthographic representations, the auditory input can influence the competition process in the visual system, and thus the decision latencies.

With these assumptions in mind, we can now discuss each of the eight RT patterns to targets preceded by voiced primes without prevoicing. We will start with the condition in which the voiced prime was a word with a voiceless word competitor (the Bear condition). When targets started with a voiced plosive (e.g., *beer)*, there was no difference between the effects of primes with prevoicing and the effects of primes without prevoicing. In both cases, decisions were faster than decisions to these targets preceded by unrelated primes. However, when targets were voiceless (e.g., *peer)*, there was a difference in the speed of lexical decisions to targets preceded by primes with prevoicing and primes without prevoicing: Primes without prevoicing resulted in faster RTs than primes with prevoicing. Nevertheless, primes without prevoicing still resulted in slower RTs than voiceless primes. We can explain these findings in the following way: Voiced primes without prevoicing activate both voiced and voiceless word candidates, but the former will be activated more strongly than the latter. Thus, upon hearing *beer* without prevoicing, the phonological representations of both *beer* and *peer* will be activated, which in turn will spread activation to the matching orthographic representations of these words. When the target is voiced, the visual input will result in strong activation of the orthographic representation of *beer*. This representation thus receives strong evidence from both the visual input and the auditory input, while the orthographic representation of *peer* is only weakly activated via the auditory system. We assume that this activation is too weak to seriously compete with *beer* under these circumstances and therefore does not influence the lexical decision on the target. When the target is voiceless, however, the orthographic representation of *peer* will receive strong activation as a result of the visual input and some extra activation via the weakly activated phonological representation of *peer.* The orthographic representation of *beer* will also be activated via the activated phonological representation of *beer* and will compete with the orthographic representation of *peer.* The moderate activation resulting from the auditory prime without prevoicing can now contribute to the competition between *beer* and *peer.* This extra activation results in faster lexical decisions to *peer* than when the prime was voiced. When prime and target are both voiceless, however, there is effectively no activation of *beer* in either the auditory or the visual system, and thus the activation of *peer* can not be suppressed by the voiced word competitor. As a result, the voiceless prime results in faster lexical decisions to the voiceless target than the voiced prime without prevoicing does.

These claims are supported by the findings of the conditions in which the voiced counterpart was a nonword, namely the Brince conditions. Under these conditions, only the voiceless lexical candidate (e.g., *prins)* could be activated (*brins* is a nonword). When the target was voiced and thus a nonword, the voiced prime without prevoicing resulted in slower “no” decisions to the nonword target than when the prime was prevoiced. Lexical decisions were even slower when the prime was voiceless. When the target was voiceless and thus a word, however, both the voiced prime without prevoicing and the voiceless prime resulted in equally faster lexical decisions in comparison to the prevoiced primes. The explanation for these findings is as follows: The voiced prime without prevoicing can now only activate the voiceless word candidate *prins,* since there is no voiced lexical candidate. Furthermore, we assume that the phonological representation of *prins* will be activated more strongly upon hearing a voiceless prime than upon hearing a voiced prime without prevoicing. As a result, more activation will spread to the orthographic representation of *prins* in the former case than in the latter. When the target is voiced and thus a nonword, the activation of the word *prins* slows down the decision that *brins* is a nonword. The difference in the degree of activation of *prins* is measurable as a difference in RTs of the “no” responses: There were slower responses to *brins* after voiceless primes than after voiced primes without prevoicing.

This difference in activation of the voiceless lexical candidate in the auditory system was, however, not detected when the target was voiceless. Under these conditions, the visual input will strongly activate the orthographic representation of voiceless words such as *prins.* The extra activation that this representation receives via the auditory system when the prime is voiceless or voiced without prevoicing will speed up lexical decisions. Although the speed of the lexical decision to the word target *prins* is influenced by the activation of the phonological representations of *prins* as a result of both types of primes, the cross-modal priming task seems to be insensitive to the small differences in the degree of activation in the auditory system in this case. The difference in phonological lexical activation can only be detected experimentally when there is conflict between the visual and auditory material. When *brins* is the visual target, such conflict does exist, and in this case the relative degree of activation of the phonological representations of *prins* does influence performance.

Exactly the same pattern was observed in the Bear conditions. When the target is voiced (e.g., *beer*) and the prime is voiced without prevoicing, there is no conflict between the visual and auditory evidence, since the visual input unambiguously supports the voiced candidate (*beer)*, and the auditory input gives more support to the voiced representations (/b/ and *beer*) than to the voiceless representations (/p/ and *peer)*. Thus *beer* can win the visual competition and no difference in the degree of activation of phonological representations as a function of the presence or absence of prevoicing is seen. When the target is voiceless (e.g., *peer*), however, there is conflict between the visual evidence (supporting "p" and *peer*) and the auditory evidence (stronger support for /b/ and *beer* than for /p/ and *peer*). So graded activation effects in the auditory system could be detected (stronger activation of *peer* by voiceless primes than by voiced primes without prevoicing). Across the Bear and Brince conditions, therefore, it appears that graded effects of prevoicing variation on lexical activation emerge when there is conflict between the visual evidence and the word most strongly activated by the auditory evidence.

This account is also supported by the results in the Blem conditions. Irrespective of the nature of the visual target (e.g., the voiced nonword *blem* or the voiceless nonword *plem*), there is no substantial activation of phonological word representations in the auditory system, and hence no conflict with the visual information. Participants found it just as easy to make "no" decisions to both of these targets when the prime had 6 periods of prevoicing as when the prevoicing was deleted. Likewise, in the case of the voiced word targets (e.g., *blauw*) in the Blue conditions, there was no effect of prevoicing deletion. Here again there is no conflict between the information delivered by the visual input and that delivered by the prime (*blauw* is the most strongly activated word; *plauw* is a nonword).

It might appear, however, that this account is challenged by the results in the eighth and final condition (the Blue condition with voiceless nonword targets such as *plauw*). Under these conditions there is conflict between the visual information (consistent with "p" and *plauw*) and the word most strongly activated by the prime (*blauw*). But, critically, there is no voiceless competitor activated by either the voiced prime with prevoicing or the voiced prime without prevoicing, since *plauw* is a nonword. The voiced candidate can thus win the competition in the auditory system in both cases, resulting in no differential priming effect. Graded activation of lexical representations as a result of prevoicing deletion thus can only be detected when a voiceless lexical competitor is activated by the input (i.e., in the Bear and Brince conditions) and even then only when there is conflict between the visual evidence and the word most strongly activated by the auditory evidence. In the absence of this conflict, competition in the visual system can resolve in favor of the target letter string.

In summary, the eight lexical decision patterns show that the extent to which graded activation in the auditory system as a result of prevoicing variation could influence the speed of lexical decisions in the cross-modal identity priming task depended on the following factors: The lexical status of the auditory prime, the voicing of the initial phoneme of the visual target, the lexical status of the target and the competitor environment of prime and target. But most importantly, the results of Experiments 4.1 and 4.2 show that the deletion of prevoicing does result in graded activation at both the prelexical and lexical level. Voiced plosives without prevoicing activate both the voiced and voiceless representations at the prelexical level. This graded activation is then passed on to the lexical level. As a result, items starting with voiced plosives without prevoicing activate both lexical candidates with voiced plosives and lexical candidates with voiceless plosives. The pattern of priming, however, then depends on the competitor environment and the nature of the visual target.

The discrimination experiment confirmed that plosives without prevoicing result in different lexical activation patterns than plosives with prevoicing. Listeners were able to discriminate among all three prevoicing durations, but their accuracy depended not only on the size of the difference (the difference between 12 and 0 periods of prevoicing was easier to detect than between 6 and 0 periods of prevoicing), but also on the presence of a token without prevoicing (the difference between 12 and 6 periods of prevoicing was harder to detect than the difference between 6 and 0 periods of prevoicing). Accuracy also depended on the lexical competitor environment, but only if one of the tokens had no prevoicing (the difference between 6 and 0 periods of prevoicing was easier to detect when there was a voiceless word competitor than when there was not). These results again indicate that tokens with plosives without prevoicing activate both voiced and voiceless plosives and, in turn, word candidates starting with voiced plosives and word candidates starting with voiceless plosives.

Overall, the three experiments provide us with sufficient evidence that differences in the duration of prevoicing (12 or 6 periods) did not influence lexical access, while the deletion of prevoicing did. Note that the pattern of results on the effects of prevoicing deletion suggest that a difference in the effect of 12 and 6 periods of prevoicing would only appear in the conditions where a voiceless word candidate could have been activated and where there is conflict between the visual evidence and the word most strongly activated by the auditory evidence. If there were graded activation of lexical candidates as a result of variation in the duration of the prevoicing (12 versus 6 periods) there ought to have been faster “no” responses to voiced nonword targets in the Brince condition when these targets were preceded by primes with 12 periods of prevoicing than when these targets were preceded by primes with 6 periods of prevoicing. Similarly, there ought to be slower “yes” responses to voiceless word targets in the Bear condition when these targets were preceded by primes with 12 periods of prevoicing than when these targets were preceded by primes with 6 periods of prevoicing. Our results suggest, however, that the patterns of activation caused by these two types of primes with prevoicing are the same. In contrast, our results show that the patterns of lexical activation for primes with prevoicing do differ from those for primes without prevoicing.

When we compare the result of the present study with those of Andruski et al. (1994), two interesting differences emerge. First, in Dutch, no graded effects were found as a result of variation in negative VOT (6 or 12 periods of prevoicing), while in English graded effects were found as a result of variation in positive VOT (unaltered or -2/3 VOT). Second, in contrast to the variation in the duration of prevoicing, the deletion of prevoicing does affect lexical activation in Dutch. Note that the equivalent case in English (i.e., deleting all aspiration and thus making the VOT zero) was not tested. Nevertheless, it is interesting to observe that the deletion of prevoicing in Dutch resulted in weaker graded effects than shortening positive VOT in English. Recall that Andruski et al. (1994) found differences in lexical decision latencies to targets preceded by related primes with unaltered VOT and -2/3 VOT, while the present priming experiments showed no difference in lexical decision latencies to targets preceded by identical voiced primes with prevoicing and without prevoicing. Crucially, however, we did find clear indications that the absence of prevoicing resulted in activation of both voiced and voiceless word candidates.

Although both studies address the same basic question, there are some important differences between the two studies which may account for the different findings. First, different experimental paradigms were used: while Andruski et al. made use of the associative priming task in one modality, we made use of the cross-modal identity priming task. One might suggest that the lexical competition process plays a more important role in the identity priming task than in the associative priming task. Therefore, the latter task may be more sensitive to graded lexical activation than the former. However, in Chapter 3 a similar pattern of results was found to the one observed here using a cross-modal associative priming task with Dutch materials. No differences were found in lexical decisions to targets like *roos* (rose) between the priming conditions in which the related prime *bloem* (flower) started with 12 or 6 periods of prevoicing and the priming condition in which the same prime had no prevoicing. All three primes resulted in the same amount of facilitation (in comparison to the unrelated prime). Furthermore, Utman et al. (2000) found graded effects of VOT variations in English using within-modality identity priming similar to those found by Andruski et al. (1994). Finally, McMurray et al. (2002) also found graded effects of English VOT variation on lexical access while using yet another paradigm (eye-tracking). This suggests that the difference between the Dutch and English findings should be sought in differences in the languages rather than in the experimental paradigms.

In order to understand the different patterns found in the English studies and the present study it is important to focus on the differences in VOT between English and Dutch. Although VOT in both languages refers to the same temporal relation, namely the moment in time that the vocal folds start vibrating in relation to the moment in time that the closure is released, acoustically and perceptually those two VOTs are very different. In English, the VOTs of voiced and voiceless plosives both fall in the positive VOT range. Voiced plosives have a small positive VOT (roughly equal to the duration of the plosive burst), while voiceless plosives have a larger positive VOT (the combined duration of the plosive burst and the following aspiration). A large amount of research has shown that English listeners are very sensitive to variation in positive VOT. A small difference in VOT can abruptly change the proportion of responses from one voicing category to the other (Liberman, Harris, Eimas, Lisker & Bastian, 1961) and the placement of the voicing boundary and the location of the best exemplars on the VOT continuum depend on speaking rate (Green & Miller, 1985; Miller, 1981; Miller & Liberman, 1979; Miller & Volaitis, 1989; Summerfield, 1981). The boundary between voiced and voiceless plosives in terms of VOT is therefore not fixed but varies on a continuous scale (see also Repp & Liberman, 1987).

In contrast to English, the main cue for the voicing distinction in Dutch is the presence or absence of voicing during the closure (van Alphen & Smits, submitted). All initial plosives with prevoicing in the van Alphen and Smits' study were perceived as being voiced by their group of listeners. The presence of prevoicing is therefore a very strong indication that the plosive is voiced in Dutch. Furthermore, that study reported two important observations for the interpretation of the results of the present study. First, the voicing percept appeared to be unaffected by the duration of prevoicing, which varied considerably within and among speakers. When a token started with prevoicing it was unambiguously perceived as voiced, regardless of the exact duration of the prevoicing. Second, in conflict with the finding that prevoicing is the primary cue to the voicing distinction, prevoicing is frequently not produced. No less than 25% of the voiced tokens in the sample of 640 utterances were produced without prevoicing. Nevertheless, the majority of these voiced tokens without prevoicing were still perceived as voiced. Other acoustic cues still signaled that the plosive was voiced, but the absence of prevoicing reduced the probability that the plosive was voiced.

These acoustic and perceptual differences between VOT in English and VOT in Dutch could serve as an explanation for the contrast between the findings of Andruski et al. (1994), Utman et al. (2000) and McMurray et al. (2002) and those of the present study. Small differences in VOT seem to be more important in English than in Dutch and the recognition system seems to treat VOT variations differently in the two languages. Since in Dutch the important cue is the presence or absence of prevoicing, and not the exact duration of prevoicing, it seems very plausible that Dutch listeners learn to ignore differences in the amount of prevoicing that is present. Keating, Mikos and Ganong (1981) have made the same suggestion about perception of initial plosives in Polish, a language which, like Dutch, makes a contrast between voiced plosives that have negative VOTs and voiceless unaspirated plosives that have positive VOTs. In contrast, as Keating et al. also suggest, English listeners should be very sensitive to gradient differences in positive VOT, since those differences carry important information for the distinction between voiced and voiceless plosives.

Furthermore, Dutch listeners might also have learned that a plosive without prevoicing could still be voiced. After all, Dutch listeners have often encountered words starting with plosives without prevoicing that should have started with voiced plosives (e.g., hearing *blauw* without prevoicing). However, since most plosives without prevoicing are actually voiceless, listeners should not ignore the presence or absence of prevoicing. The results of the current study show that listeners are indeed sensitive to this difference. When prevoicing is absent, both voiced and voiceless prelexical representations are activated, and the voiced prelexical representation is activated more strongly upon hearing a word with prevoicing than upon hearing a word without prevoicing. As a result both voiced and voiceless word candidates are activated. In contrast, there is no similar kind of ambiguity in VOT in English natural speech. English listeners will have rarely if ever encountered words starting with plosives without a positive VOT that should have been voiceless according to the lexicon. It would therefore be absurd if English listeners were to treat plosives with zero VOT as voiceless.

Both the English and Dutch recognition systems thus seem to be sensitive to VOT variation that provides important information for the distinction between voiced and voiceless plosives. In Dutch the presence or absence of prevoicing is important, while in English the duration of the positive VOT is important. In line with this, the present study shows that, in Dutch, lexical access is influenced by the presence or absence of prevoicing, while Andruski et al. (1994), Utman et al. (2000) and McMurray and Tanenhaus (2002) found that, in English, lexical access is influenced by variations in positive VOT. Furthermore, as we noted earlier, the effect of prevoicing deletion on lexical access in Dutch seems to be less strong than the effect of shortening the positive VOT in English. This can be explained by the finding that in Dutch prevoicing is frequently absent (van Alphen & Smits, submitted).

So far we have mainly concentrated on the differences between the effects of VOT variation in Dutch and English speech recognition. The comparison between the two languages indicate that the recognition system has adapted itself to the speech input in that particular language by taking into account the informational value and the natural distributions of particular acoustic properties. As a result of this, different graded effects are found for VOT variations in Dutch and English. Despite these differences, however, the findings of graded effects in both languages lead to the same claims about the general mechanisms of the recognition system, which are shared by current models of word recognition such as TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994; Norris et al., 2000).

First, the results support the assumption that multiple lexical candidates are activated in parallel and that the initial degree of activation of these candidates reflects the goodness of fit between the incoming speech signal and lexical representations. Second, the results indicate that lexical access is continuous; the prelexical processor passes information to the lexical level in cascade. If the prelexical processor acted in a serial and categorical fashion, there ought to be no effects of continuous phonetic variables on lexical access. Models in which an absolute phonemic categorization of the input is made prior to lexical access are therefore inconsistent with the present results. These results, however, are not informative about the size of units at the prelexical level. Although we have explained some of the data in this article by using the idea of prelexical representations of voiced and voiceless plosives, we do not want to claim that the representations at the prelexical level are necessarily phonemic. Any smaller units (e.g., allophones or features) or larger units (e.g., diphones or syllables) would be consistent with the present findings, so long as those units pass activation in cascade up to the lexical level. Third, there is competition between activated lexical candidates. The present study demonstrates the large impact that the competition process has on word recognition.

CONCLUSION

We have shown that Dutch listeners are sensitive to variations in prevoicing in Dutch initial voiced plosives, but only when prevoicing varies from being present to being absent. Variations in the duration of prevoicing in the cases were there is prevoicing seems not to influence the activation of representations at the prelexical or lexical level. Both 12 and 6 periods of prevoicing appear to be good exemplars of the voiced category. As a result, words starting with voiced plosives with either 12 or 6 periods activate voiced lexical candidates to the same degree. The prelexical processor, however, does appear to be sensitive to the deletion of prevoicing. Deleting the prevoicing of voiced plosives takes away an important cue to the voicing of the plosive, resulting in somewhat ambiguous plosives. These plosives are in general still perceived as being voiced. Upon hearing an item starting with a voiced plosive without prevoicing, both voiced and voiceless representations at the prelexical level are activated.

The differences between these two types of VOT variations (variation in the duration of prevoicing, and the presence or absence of prevoicing) can be explained in terms of the informational value of these variations. The prelexical processor appears to be only sensitive to variation in the speech signal that is informative (6 vs. 0 periods of prevoicing). Variation in the exact duration of prevoicing (12 vs. 6 periods of prevoicing) is uninformative in Dutch and is therefore normalized away at the prelexical level. As a result, only the difference between the presence and absence of prevoicing influences lexical access. We have therefore shown that while phonetic fine detail can influence lexical processing in a continuous and graded fashion, such effects are conditional on whether the fine detail is relevant for lexical distinctions and thus has communicational value. The comparison between Dutch and English indicates in addition that variation in VOT across languages appears to influence word recognition as a function of the informational value of VOT for lexical processing in a particular language.

SUMMARY AND CONCLUSIONS

CHAPTER 5

One interesting and essential component of speech recognition is the way in which the listener maps the highly variable acoustic signal onto stored lexical representations in the mental lexicon. There is no simple key which the listener can use to crack the speech code, since there are no one-to-one relationships between acoustic properties and stored mental representations. Nevertheless, listeners seem to have no problem at all extracting the relevant cues from the speech signal and ignoring the acoustic variation that is irrelevant for the recognition of words. This thesis systematically investigated the influence of relevant versus irrelevant acoustic detail on lexical processing and the way in which these effects were influenced by the existence of strong lexical competitors. The acoustic property which was examined was prevoicing. Prevoicing refers to negative Voice Onset Time. It is said to be one of the primary cues to the voicing distinction in Dutch initial plosives.

The purpose of Chapter 2 was to establish the range of prevoicing variation in natural Dutch speech. In order to be able to make predictions about the effects of prevoicing variation on lexical access, it was important to quantify how prevoicing varies. In the first experiment of this chapter, 10 Dutch speakers were asked to produce 64 items starting with a voiced plosive. The results showed that no less than 25% of these tokens were produced without prevoicing. Furthermore, the results showed that the duration of prevoicing, when it was present, varied considerably. Several factors appeared to have an influence on the proportion of prevoiced tokens and on the duration of the prevoicing that was present. Male speakers produced more voiced tokens with prevoicing than female speakers did. This difference is probably caused by a difference in vocal tract size between males and females. A larger vocal tract size will make it easier to initiate and maintain vocal fold vibration during the closure of the plosive. The difference in the duration of prevoicing between male and female speakers was however not significant. Furthermore, there were two segmental factors which affected the proportion of prevoicing, namely the place of articulation of the plosive and the type of phoneme which followed the plosive. Labials (i.e., /b/s) were more often produced with prevoicing than alveolars (i.e., /d/s), and plosives followed by a vowel were more often prevoiced than plosives followed by a consonant. The latter factor also had a significant effect on the duration of prevoicing. These two segmental effects can be explained in terms of the possible degree of vocal tract expansion during plosive closure. Expansion of the vocal tract, passively or actively, makes it easier to produce prevoicing since the supraglottal pressure rises less quickly. The more anterior constriction in the case of labial plosives in comparison to alveolar plosives results in a larger area of oral cavity surface which can yield to passive expansion. In the case of alveolar plosives, the pharyngeal walls and part of the soft palate can be pushed outward in reaction to the raised pressure, while during the production of a labial plosive these surfaces plus parts of the cheek can participate in the expansion. The effect of the following phoneme is probably also due to differences in the freedom to expand the vocal tract volume, although the exact mechanisms of expansion in each of the different contexts which were used are less clear. The two lexical factors which were examined, namely the lexical status of the carrier word and the presence of a voiceless word competitor, affected neither the proportion of prevoicing nor the prevoicing duration.

It thus seems that prevoicing is most frequently absent in those circumstances in which the aerodynamics make it more difficult to produce prevoicing. Although this might suggest that in some cases the production of prevoicing is simply too difficult, this explanation is not satisfying, since studies on other languages, for example on Polish (Keating, Mikos & Ganong, 1981) or Lebanese Arabic (Yeni-Komshian, Caramazza & Preston, 1977) did not report such a high proportion of unprevoiced tokens. It could be the case, however, that the frequent absence of prevoicing is a result of the influence of the English language on Dutch speakers. Caramazza and Yeni-Komshian (1974) found that in Canadian French prevoicing is frequently omitted, while in European French prevoicing is rarely omitted. They argued that these observed changes in Canadian French were a result of the frequent contact of the Canadian French speakers with English. A similar argument could be made for Dutch. The frequent exposure of Dutch speakers to the English language, in which voiced plosives are in general produced without prevoicing, may be an explanation for the observation that prevoicing is frequently absent in Dutch.

Note, however, that the 25% unprevoiced tokens were all produced at the beginning of an utterance. In natural speech, voiced plosives frequently appear in the middle of utterances. When the phoneme preceding the voiced plosive is voiced, the production of prevoicing might be slightly easier, since in that case the vocal cords are already vibrating. It is therefore more likely that at least part of the closure is voiced. On the other hand, the proportion of 25% is based on speech produced under experimental conditions. I assume that during natural conversation speakers are less careful in their production; prevoicing may therefore be absent even more frequently, at least at the beginning of an utterance or after voiceless phonemes. One way in which a more precise estimation of the proportion of unprevoiced voiced plosives in Dutch could be obtained would be to measure the prevoicing of voiced plosives in a corpus of spontaneous speech. Such a corpus (Corpus Gesproken Nederlands) is currently being assembled, but is not yet available.

It is well known that distinctions between phonemes, and thus between words, are signalled by multiple cues, some being more important than others. The finding that no less than 25% of the voiced plosives in the analysis presented in Chapter 2 were produced without prevoicing immediately raises the question how these tokens are perceived and whether prevoicing is in fact such an important cue. Therefore, the second experiment presented in Chapter 2 examined the influence of several potential acoustic cues, including prevoicing, on the perception of the voicing distinction. The following measures were obtained from a sample consisting of 48 voiced-voiceless pairs produced by 10 speakers: prevoicing duration, burst duration, power of the burst above 500 Hz, spectral centre of gravity of the burst, F_0 immediately after the burst and the F_0 difference between the F_0 immediately after the burst and the F_0 in the middle of the (consonant plus) vowel. All these measures, except the F_0 immediately after the burst, showed a significant difference between voiced and voiceless plosives. A CART analysis predicted that the presence or absence of prevoicing would be the best predictor for the voicing distinction. A CART analysis on the proportion of voiced responses, which were derived from the identification responses given by 16 listeners, showed that the listeners indeed relied most strongly on prevoicing when identifying the tokens as voiced or voiceless. Practically all tokens produced with prevoicing were perceived as voiced and the majority of the tokens without prevoicing were perceived as voiceless. Nevertheless, a proportion of the tokens without prevoicing were still perceived as voiced on the basis of other acoustic cues. The relative strength of these acoustic cues differed for the two places of articulation: for the labial /b/, the F_0 difference appeared to be the strongest secondary cue, while for the alveolar /d/, the spectral centre of gravity appeared to be the strongest secondary cue.

The finding that prevoicing is the strongest perceptual cue for voiced plosives, even though prevoicing is frequently absent, makes the question about the effects of prevoicing variation on lexical activation even more interesting. On one hand one might predict that the deletion of prevoicing would have a strong effect on lexical activation, since it would take away a very strong indication that the plosive was voiced. In the majority of the cases, plosives without prevoicing are after all intended to be voiceless. On the other hand, prevoicing is frequently absent in naturally produced voiced plosives. Listeners thus have often encountered words in their life with an initial voiced plosive that have been produced without prevoicing. It might therefore have been the case that the absence of prevoicing does not strongly affect lexical access. The effects of prevoicing deletion on lexical processing were investigated in Chapters 3 and 4.

Before summarizing the findings of these experiments, I would like to briefly compare the items without prevoicing in the identification experiment of Chapter 2 with the items without prevoicing used in the priming experiments in Chapters 3 and 4. The difference between those two groups of items is that the former group of items without prevoicing were naturally produced without prevoicing while the latter group of items were obtained by deleting prevoicing that was present in the initial recording. Without doing any acoustic measurements on the manipulated items, it is hard to predict what the difference between these two groups of items would be, if there were any differences. It is possible that the secondary acoustic cues of voiced plosives which are naturally produced without prevoicing are stronger than the secondary acoustic cues of voiced plosives from which the prevoicing was deleted. That is, one could argue that speakers try to compensate for the lack of prevoicing and therefore make the other cues stronger. Nevertheless, one could easily predict the opposite. Acoustic properties are after all not produced independently from each other; one articulatory gesture can result in multiple acoustic cues, and therefore many cues are highly correlated. It is thus also possible that the secondary cues were stronger in the plosives which were naturally produced with prevoicing.

One way to find out which of these two possibilities is most likely to be true is to compare the mean values of the acoustic properties which were measured in the second experiment of Chapter 2 for voiced tokens produced with prevoicing and voiced tokens without prevoicing. In such an analysis, it appeared that the mean values of all five measures in the group of voiced plosives without prevoicing lay closer to the mean values of the voiceless plosives than the means in the group of voiced plosives with prevoicing. That is, voiced plosives without prevoicing showed longer burst durations with higher power values and higher spectral centre of gravity values, higher F_0 values immediately after the burst, and smaller F_0 differences. A subject-multivariate analysis of variance showed that for four measures (i.e., all measures except for the F_0 difference) the voiced plosives with prevoicing were significantly different from those without prevoicing. Thus, the plosives without prevoicing tended to have weaker secondary cues favoring the voiced alternative. This suggests that the secondary cues in the items used in the priming experiments (those items were originally produced with prevoicing) would tend to be stronger than the secondary cues in voiced tokens which are naturally produced without prevoicing. The effects of prevoicing deletion might thus have been stronger if natural unprevoiced tokens had been used. The only way to examine systematically the effects of prevoicing variation alone, however, was to use the same naturally produced token of each item to construct the different prevoicing conditions in the experiments. In this way the strength of the secondary cues was kept constant within items. It would have been impossible to control for this by using naturally produced items with and without prevoicing.

In Chapters 3 and 4 the effects of prevoicing variation on lexical processing were investigated. Two types of prevoicing differences were contrasted, namely the difference between the presence and absence of prevoicing (6 versus 0 periods of prevoicing) and the difference in the amount of prevoicing (12 versus 6 periods of prevoicing). Both types of differences were of the same size, namely six periods of prevoicing, and fell within the natural range of prevoicing variation. There was however a difference in the perceptual relevance of the two types of prevoicing variation. While the presence or absence of prevoicing is a strong cue to the voicing distinction (as we have seen in Chapter 2), the exact duration of the prevoicing which is present is not. The prediction was therefore that the difference between 6 and 0 periods of prevoicing would result in different degrees of lexical activation, while the difference between 12 and 6 periods of prevoicing would be normalized away at an earlier level of processing.

In Chapter 3 two priming experiments were presented in which the effects of the two types of prevoicing differences were investigated. The first experiment was a cross-modal associative priming experiment. Primes were either high or low frequency words starting with voiced plosives and were either semantically related to the target, for example *bloem* (flower) and *roos* (rose), or semantically unrelated, for example *baan* (job) and *roos* (rose). Each related prime appeared in three different prevoicing conditions: with 12, 6 or 0 periods of prevoicing. The results showed a clear priming effect of the related primes, that is, lexical decisions to targets were faster when preceded by related primes than when preceded by unrelated primes. There was however no difference among the three prevoicing conditions. The same materials were then used in a cross-modal identity priming task. In addition to the word primes, there were now also nonword primes. Primes were now either identical to the target, e.g. *bloem* and *bloem*, or unrelated, e.g. *baan* and *bloem*. The result showed again a clear priming effect for the identical primes, but there was again no effect of prevoicing variation. Although no effect of prevoicing duration was predicted in the case of the difference between 12 and 6 periods of prevoicing, the absence of an effect of prevoicing deletion was unexpected. A discrimination experiment showed that listeners were in fact able to hear the differences among all three prevoicing conditions. Thus the absence of an effect in the priming studies was not due to an inability of the listeners to hear the differences in prevoicing duration.

The finding that primes starting with voiced plosives without prevoicing resulted in similar degrees of priming as those starting with voiced plosives with prevoicing could be a result of the fact that prevoicing is frequently absent in Dutch. Thus although listeners rely strongly on the presence of prevoicing, they have also learned that when prevoicing is absent this does not always imply that the plosive is voiceless. While the presence of prevoicing clearly signals a voiced plosive, the absence of prevoicing signals that the plosive is either voiced or voiceless. If this is indeed true, one would predict that tokens without prevoicing activate both voiced and voiceless prelexical representations to a significant degree and that in turn lexical candidates starting with both voiced and voiceless plosives are activated. Nevertheless, the voiced tokens without prevoicing were not completely ambiguous. Since all other acoustic cues were preserved, these cues presumably still signalled that the voiced plosives without prevoicing were voiced. This was confirmed by the results of the identification phase which followed each priming experiment. Although the proportion of voiced responses was lower for the voiced tokens without prevoicing than the voiced tokens with prevoicing, the latter were in general still perceived as voiced. It thus appeared that the experiments presented in Chapter 3 were not sensitive to effects of prevoicing deletion. This may have been because all items used in these experiments had no voiceless word competitors. Upon hearing for example *bloem* without prevoicing, the lexical representation of *bloem* was the only serious lexical candidate, since *ploem* does not exist. Therefore, the activation of the lexical representation of *bloem*, though temporarily weaker as a result of the absent prevoicing, could quickly recover since there were no serious lexical competitors.

In Chapter 4 the influence of the presence of voiceless lexical candidates was investigated in two cross-modal identity priming experiments. Primes were words or nonwords starting with a voiced plosive which had either a voiceless word competitor or not. This resulted in four priming conditions, which were labeled as the Blue condition (word, no competitor), the Bear condition (word, with competitor), the Blem condition (nonword, no competitor) and the Brince condition (nonword, with competitor). There were again three prevoicing conditions (12, 6 and 0 periods) and an unrelated condition. This time there was also a voiceless priming condition, in which the voiceless counterparts of the voiced primes were presented. In Experiment 4.1, the voiced counterparts were used as targets (e.g., *blauw* (blue), *beer* (bear), *blem* and *brins*), while in Experiment 4.2, the voiceless counterparts served as targets (e.g., *plauw, peer* (pear), *plem* and *prins* (prince)). All conditions showed facilitation of the lexical decisions to targets which were preceded by identical primes in comparison to targets preceded by unrelated primes. When prime and target differed only in the first phoneme, for example *peer* and *beer* (or the other way around), there was no facilitation.

Interestingly, in the conditions in which the voiced item had a voiceless word competitor (the Bear condition and the Brince condition) there was an effect of prevoicing deletion. These results indicated that the primes without prevoicing had also activated the voiceless word competitors. For example, the prime *beer* without prevoicing had activated both *beer* and *peer.* The activation of the latter was however weaker than the activation of the former, since a voiced plosive without prevoicing is still more voiced than voiceless. In the case of *brins* without prevoicing, *prins* was the only serious lexical candidate, since *brins* is a nonword. However, *prins* was less strongly activated upon hearing *brins* without prevoicing than upon hearing *prins*. Thus items starting with voiced plosives without prevoicing resulted in graded activation of voiceless word competitors. This graded activation of the voiceless lexical representations could however only be detected when there was conflict between the visual evidence provided by the target and the word most strongly activated by the auditory prime. In the absence of this conflict, competition in the visual system could resolve in favor of the target letter string. For example, the graded activation of the lexical representation of *peer* could only be detected when the target was *peer*. There is conflict in this situation because, upon hearing *beer* without prevoicing, *beer* is the most strongly activated candidate. When the target was *beer*, however, there was no conflict: the degree of facilitation by *beer* without prevoicing was the same as the degree of facilitation by *beer* with prevoicing. In the case of *brins* without prevoicing, *prins* was now the most strongly activated candidate (since *brins* is a nonword), thus the graded activation of *prins* could only be detected when the target was *brins*. When the target was *prins*, the prime *brins* without prevoicing resulted in the same degree of facilitation as the voiceless prime *prins*.

These results indicate that lexical activation is affected by relevant acoustic variation but not by variation which is irrelevant for lexical distinctions. In none of the conditions was there a difference between the priming condition with 12 periods of prevoicing and the condition with 6 periods of prevoicing. Both prevoicing durations resulted in similar lexical activation patterns. In contrast, the difference between 0 and 6 periods of prevoicing did affect lexical activation, but only in the conditions where there was a voiceless lexical candidate which could have been activated.

These findings were strengthened by the results of a discrimination experiment. In this experiment, listeners heard all six prevoicing combinations for each item. They were asked to indicate whether the initial sound of the two members of a pair were the same or different. In line with the findings of the discrimination experiment presented in Chapter 3, the results showed that it was easiest to discriminate between 12 and 0 periods of prevoicing, more difficult to discriminate between 6 and 0 periods and most difficult to discriminate between 12 and 6 periods of prevoicing. Furthermore, there was a clear effect of the presence of a voiceless lexical competitor: it was easier to discriminate between two members with different degrees of prevoicing when these members had a voiceless word competitor, but only when one of the members had no prevoicing. These findings again suggest that items starting with voiced plosives without prevoicing also significantly activate voiceless lexical candidates. Therefore, the difference between 6 and 0 periods of prevoicing affected lexical access, while the difference between 12 and 6 periods of prevoicing did not.

Reading the summaries of the preceding chapters an observant reader might note that while in the acoustical-perceptual studies a distinction was made between labial and alveolar plosives, no such distinction was made in the priming or discrimination studies. The results of Chapter 2 do however suggest that the role of prevoicing is different for the two types of plosives. Labial plosives appeared to suffer more from the absence of prevoicing than alveolar plosives did. No less than 52% of the voiced labials without prevoicing were perceived as voiceless, while this proportion was only 26% for the voiced alveolars without prevoicing. Therefore, the other acoustic cues to the voicing distinction seemed to be stronger in the alveolar plosives. The presence of prevoicing thus appears to be less important in the perception of the voicing distinction in alveolar plosives than in labial plosives.

This difference in the importance of prevoicing between labial and alveolar plosives was however not examined in the priming experiments of Chapters 3 and 4. On the basis of the results of the acoustical-perceptual experiments of Chapter 2 one would predict that the effects of prevoicing deletion would be stronger for items starting with labial voiced plosives than for items starting with alveolar voiced plosives. In other words, one would expect that upon hearing an item starting with a /b/ without prevoicing a possible voiceless lexical candidate (e.g., a /p/-initial word) would be activated more strongly than /t/-initial words might be activated upon hearing a /d/-initial item without prevoicing. Unfortunately, there were not enough items in each of the conditions to investigate the effect of place of articulation.

In summary, the experiments presented in Chapters 3 and 4 indicate that lexical activation is sensitive to the difference between 6 and 0 periods of prevoicing, but not to the difference between 12 and 6 periods of prevoicing. This is in line with the prediction on the basis of the relevance of these two types of prevoicing differences. The difference between the presence and absence of prevoicing is relevant for lexical distinctions, since it is an important cue to the voicing distinction in Dutch. The amount of prevoicing, when present, however, is not informative, since the presence of prevoicing clearly signals that a plosive is voiced, regardless of its duration. As a result, only the former type of prevoicing variation can affect lexical activation. Interestingly, prevoicing is frequently absent in natural speech. Therefore, listeners can not fully rely on prevoicing. Although the presence of prevoicing unambiguously signals that the plosive is voiced, the absence of prevoicing does not unambiguously signal that it is voiceless. Listeners must have learned that when there is no prevoicing the plosive could be either voiced or voiceless. Note, however, that deleting the prevoicing of a voiced plosive does not result in a fully ambiguous token, since all the other acoustic cues are preserved and thus still signal that the plosive is voiced. Nevertheless, when a voiced plosive does not have prevoicing, the probability that the plosive is voiced is decreased. As a result, voiced plosives without prevoicing activate voiced prelexical representations and voiceless prelexical representations, giving more support to the voiced than the voiceless one. This graded activation is passed on continuously to the lexical level.

It then depends on the lexical neighborhood what will happen. When there is only a voiced lexical candidate which closely matches the acoustic input, it can easily win the competition, since there are no other serious competing candidates present. In this case, the absence of prevoicing seems not to affect word recognition. When there is also a voiceless candidate, however, both the voiced and voiceless lexical candidates are activated. These two candidates will then compete with each other and in the end the voiced candidate will win since it forms a slightly better match to the acoustic input than the voiceless candidate. Lexical activation is thus sensitive to the absence of prevoicing, but it does not hinder the recognition of words starting with a voiced plosive. Listeners apparently use the knowledge they have about the way in which prevoicing varies in natural speech; they know that prevoicing can help them in the matching process between acoustic input and stored lexical representations, but they also know that it is not fully reliable.

What do these results tell us about the way in which the prelexical processor operates? First, the finding that the difference between the presence and absence of prevoicing resulted in different lexical activation patterns, while the difference in the duration of prevoicing did not, indicates that relevant acoustic differences are preserved by the prelexical level, while irrelevant information is normalized away. This could be accounted for by a model in which the prelexical units represent phonemes. In that case, variation in the amount of prevoicing results in the same degree of activation of the voiced phoneme, for example [b], while differences in the presence or absence of prevoicing result in different degrees of activation of not only the voiced phoneme, but also the voiceless phoneme, for example [p]. Upon hearing a labial plosive with prevoicing, [b] is strongly activated, since the presence of prevoicing is a strong indication that the plosive is voiced. Upon hearing a labial plosive without prevoicing, however, [b] is less strongly activated than upon hearing a prevoiced token, and in addition [p] is also significantly activated, but less strongly than [b]. The same story would hold if one were to assume that the prelexical units represent phonological features rather than phonemes (for example [+voice] and [-voice]). Crucially, however, models with phonemes or phonological features at the prelexical level can only account for the data of the present experiments if it is assumed that these prelexical units are activated in a graded fashion and that this graded activation is passed on to the lexical level. A model such as FUL (Lahiri & Reetz, 1999), in which a phoneme is either [+voice] or [-voice] and in which no graded activation of these features is allowed, is proven to be incorrect by the present findings. The original implementation of Shortlist (Norris, 1994) in which the input to the lexical level consisted of a string of phonemes, could therefore also not account for the current findings. A more recent version of Shortlist (Norris, McQueen & Cutler, 2000) does involve graded activation of prelexical representations. Furthermore, the model is currently being improved through addition of probabilistic information about the incoming speech signal. In this way the input to the lexical level will consist of a string of time-varying phoneme probabilities. These probabilities were derived from the results of a large gating experiment involving all 1,179 diphones of Dutch, gated at six points during each diphone (Smits, Warner, McQueen & Cutler, 2003). In this way, the input to the lexical level is similar to the continuous output of a prelexical level at which phonemes are activated in graded fashion.

The results presented in this thesis could however also be explained by models (e.g., the LAFF model, Stevens, 2002) in which the prelexical units are smaller than phonemes or phonological features, for example units representing individual acoustic features. As mentioned several times in this thesis, several acoustic cues contribute to the perception of a plosive as being phonologically voiced or voiceless. Therefore, it is possible that lexical representations are not linked to single phoneme representations at the prelexical level, but to a combination of representations of different acoustic features. Prevoicing would then be one of the acoustic feature representations which could help to distinguish between *beer* and *peer,* but prevoicing would not be the only one. Other features, for example the F_0 movement into the vowel, would then also be represented independently at the prelexical level. The findings that in Dutch it is the presence or absence of prevoicing rather than the degree of prevoicing that is important in word recognition might suggest that the activation of these acoustic feature representations is binary. Although this might work for prevoicing, it may not work for most other acoustic cues, for example vowel duration or VOT in English. There is no simple binary distinction in these cues which helps to distinguish among different lexical candidates. It is therefore very unlikely that the prelexical level, even if it consists of separate acoustic feature representations, makes discrete decisions.

One could however also argue that there is no prelexical level at all and that the speech signal is directly mapped onto lexical representations. These lexical representations could consist of prototypes of the acoustic realizations of each word (Klatt, 1979, 1989) or of the exemplars of all tokens of a word ever heard (Goldinger, 1998). In the latter case, the lexical representation of *beer* would contain tokens with 0, 6 or 12 periods of prevoicing (amongst other durations), while the lexical representation of *peer* would contain tokens with only 0 periods of prevoicing. The distributional properties of each acoustic property would thus be stored in the lexical representations, and thus also the knowledge about relevant and irrelevant acoustic detail. In this way, the effects which were found in the present priming experiments could also be explained by a model without a prelexical level.

Although there is no experimental evidence which conclusively proves the existence of a prelexical level, there are many arguments in favor of it. For example, the existence of prelexical representations would reduce a lot of the redundancy which would otherwise exist in the lexicon. In the absence of a prelexical level, knowledge about the phonetic properties of each speech sound has to be stored multiple times at the lexical level. The prelexical level is therefore a stage at which different acoustic properties can be normalized. As pointed out earlier, speech is highly variable and many different factors influence the exact acoustic realization of any particular speech sound. For example, there is a lot of variation among speakers due to differences in the sex and age of the speaker, the size and shape of the vocal tract, the dialect of the speaker, the speaking rate and the speech style. The variation caused by these factors are not relevant for lexical distinctions. At a prelexical level of processing the influence of these factors could be normalized away such that only the relevant acoustic variation remains. For example, in the case of English VOT, the influence of speaking rate on the VOT value could be normalized away at the prelexical level, such that only the relative VOT value would remain. In this way, the lexical level would not have to store each word produced at every possible speaking rate. It is therefore very plausible to assume that there is a prelexical level of processing.

The findings of the experiments presented in this thesis may not answer the question about the size of the prelexical units, but they show that a particular type of variation does not affect lexical activation, while another type of variation (even within the same acoustic cue) does affect the activation of lexical candidates. These differences can be explained by differences in the informational value of the two types of variation. Only relevant acoustic detail affects the goodness-of-fit between the speech signal and the stored lexical representations and thus the degree of lexical activation. The relevance of different types of acoustic detail largely depends on the way in which it occurs in natural speech. Listeners appear to be very sensitive to this type of distributional information. Furthermore, the data show an intriguing paradox in the role of prevoicing: prevoicing appears to be the primary cue to the voicing distinction in Dutch, but at the same time, prevoicing is frequently absent in voiced plosives. This paradox is reflected in the way in which the speech recognition system treats words starting with voiced plosives without prevoicing. When prevoicing is absent, not only is the lexical candidate starting with a voiced plosive activated, but also the voiceless word competitor (should there be one). Nevertheless, the deletion of prevoicing does not prevent the listener from correctly recognizing the word starting with the voiced plosive.

The way in which Michel de Montaigne described speech more than four centuries ago is thus very apt. *La parole est moitié à celuy qui parle, moitié à celuy qui l'escoute*... Speech indeed belongs to both the one who speaks and the one who listens. *Cettuy cy se doibt preparer à la recevoir, selon le bransle qu'elle prend*. The way that a listener must "prepare" in order to be able to understand the message of the speaker is through being observant to the "motions" that speech takes. By doing so, the listener learns which parts of the motions are meaningful and which are not and which motion patterns are more frequent than others. This information will help the listener to extract the message from the complex signal which is produced by the speaker.

REFERENCES

Allen, J. S. & Miller, J. L. (2001). Contextual influences on the internal structure of phonetic categories: a distinction between lexical status and speaking rate. *Perception & Psychophysics,* **63**, 798-810.

Alphen, P. M. van & McQueen, J. M. (submitted). The effect of Voice Onset Time differences on lexical access in Dutch. Manuscript submitted for publication.

Alphen, P. M. van & Smits, R. (submitted). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives; The role of prevoicing. Manuscript submitted for publication.

Andruski, J. E., Blumstein, S. E. & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition*, **52**, 163-187.

Baayen, H., Piepenbrock, R.,& Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Balota, D. A. & Chumbley, J. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 340-357.

Berg, J. van den (1958). Myoelastic theory of voice production. *Journal of Speech and Hearing Research*, **1**, 277-244.

Booij, G. E. (1981). *Generatieve fonologie van het Nederlands.* Utrecht: Het Spectrum.

Booij, G. E. (1995). *The phonology of Dutch.* Oxford: Oxford University Press.

Breiman, l., Friedman, J. H., Olshen, R. H. & Stone, C. J. (1984). *Classification and regression trees.* New York: Chapman & Hall.

Caramazza, A. & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics,* **2**, 239-245.

Cho, T. & Ladefoged, P. (1999). Variations and universals in VOT: Evidence from 18 languages. *Journal of Phonetics,* **27**, 207-229.

Cho, T., Jun, S. & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics,* **30**, 193-228.

Cohen, A., Ebeling, C. L., Fokkema, K. & van Holk, A. G. F. (1969). *Fonologie van het Nederlands en het Fries.* ‘s-Gravenhage: Martinus Nijhoff.

Connine, C. M., Blasko, D. G. & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, **32**, 193-210.

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

Cutler, A., van Ooijen, B. & Norris, D. (1999). Vowels, consonants and lexical activation. In J.J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A.C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences*, Vol. 3 (pp. 2053-2056). Berkeley: UC Berkeley.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K. & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, **16**, 507-534.

Davis, M. H., Marslen-Wilson, W. D. & Gaskell, M. G. (2002). Leading up the lexical garden- path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 218-244.

Ernestus, M. (2000). *Voice Assimilation and segment reduction in casual Dutch*. Utrecht: Holland Institute of Generative Linguistics.

Flege, J. E. and Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, **6**, 185-202.

Forrest, K., Weismer, G., Milenkovic, P. & Dougall, R. N. (1988). Statistical-analysis of word- initial voiceless obstruents - Preliminary data. *Journal of the Acoustical Society of America* **84**, 115-123.

Foss, D. J. & Blank, M. A. (1980). Identifying the speech codes. Cognitive Psychology, **12**, 1-31.

Frauenfelder, U. H., Scholten, M. & Content, A. (2001). Bottom-up inhibition in lexical selection: Phonological mismatch effects in spoken word recognition. *Language and Cognitive Processes*, **16**, 583-607.

Garnsey, S. M., Pearlmutter, N. J., Myers, E. & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language,* **37**, 58-93.

Goldinger, S. D. (1998). *Echoes of echoes?: An episodic theory of lexical access.* Psychological Review, **105**, 251-279.

Gow, D. W. (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 163-179.

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

Grainger, J. & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review,* **103**, 518-565.

Green, K. P. & Miller, J. L. (1985). On the role of visual rate information in phonetic perception. *Perception & Psychophysics*, **38**, 269-276.

Haggard, M. P., Ambler, S. & Callow, M.(1970). Pitch as a voicing cue. *Journal of the Acoustical Society of America*, **47**, 613-617.

Haggard, M., Summerfield, Q. & Roberts, M. (1981). Psycho-acoustical and cultural determinants of phoneme boundaries - Evidence from trading F0 cues in the voiced- voiceless distinction. *Journal of Phonetics,* **9**, 49-62.

Honda, K., Hirai, H. & Kusakawa, N. (1993). Modeling vocal tract organs based on MRI and EMG observations and its implication on brain function. *Annual Bulletin, No. 27, research Institute of Logopedics and Phoniatrics*, University of Tokyo, 37-49.

Houde, R. A. (1968). A study of tongue body motion during selected speech sounds. *Speech Communication Research Labatory,* Santa Barbara, Monograph No. 2.

House, A. S. & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America,* **25**, 105-113.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language,* **60**, 286-319.

Keating, P. A., Linker, W. & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics,* **11** , 277-290.

Keating, P. A., Mikos, M. J. & Ganong, W. F. (1981). A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America,* **70**, 1261-1271.

Keppel, G. (1982). *Design and analysis: A researcher's handbook.* Englewood Cliffs, NJ: Prentice Hall, Inc.

Kewley-Port, D. & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, **2**, 195-210.

Kingston, J. & Diehl, R. L. (1994). Phonetic knowledge. *Language* **70**, 419-454.

Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, **7**, 279-312.

Klatt, D. H. (1989). Review of selected models of speech perception. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169-226). Cambridge, MA: MIT Press.

Kohler, K.J. (1985). F_0 in the perception of lenis and fortis plosives. *Journal of the Acoustical Society of America*, **78**, 21-32.

Konefal, J.A. & Fokes, J. (1981). Voice onset time: The development of Spanish-English distinction in normal and language disordered children. *Journal of Phonetics,* **9**, 437-444.

Lahiri, A. & Reetz, H. (2002). Underspecified Recognition. In C. Gussenhoven, N. Warner (Eds.), *Labphon VII* (pp. 637-675), Berlin: Mouton de Gruyter.

Lehiste, I. & Peterson, G. E. (1961). Some basic considerations on the analysis of intonation. *Journal of the Acoustical Society of America,* **33**, 419-425.

Liberman, A. M., Harris, K. S., Eimas, P. & Bastian, J. (1961). An effect of learning on speech perception: The discrimination of durations of silence with and without phonemic significance. *Language & Speech*, ***4***, 175-195.

Lisker, L. and Abramson, A. S. (1964). A cross-language study of voicing in initial stops - acoustical measurements. *Word-Journal of the International Linguistic Association* **20**, 384- 422.

Lisker, L. & Abramson, A. S. (1970). The voicing dimension: Some experiments in comparative phonetics. *Proceedings of the Sixth International Congress of Phonetic Sciences* (pp. 563- 567). Prague: Academia.

Löfqvist, A. (1978). Intrinsic and extrinsic F0 variation in Swedish tonal accents. *Phonetica*, **31**, 228-247.

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

Lugt, A., van der (1999). *From speech to words*. Doctoral Dissertation, University of Nijmegen (MPI Series in Psycholinguistics, **13**).

Lugt, A., van der (2001). The use of sequential probabilities in the segmentation of speech. *Perception & Psychophysics*, **63**, 811-823.

Macken, M. A. & Barton, D. (1980). The acquisition of the voicing contrast in English: a study of voice onset time in word-initial stop consonants. *Journal of Child Language***,** **7**, 41-74.

Macmillan, N. A. & Creelman, C. D. (1991). Detection Theory: A User's Guide. Cambridge: Cambridge University Press.

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

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

Marslen-Wilson, W. & Zwitserlood, P. (1989). Accessing spoken words: On the importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 576-585.

Marslen-Wilson, W., Moss, H. E. & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 1376-1392.

Massaro, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry.* Hillsdale, NJ: Erlbaum.

McClelland, J. L. & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **10**, 1-86.

McMurray, B., Tanenhaus, M. K. & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, **86**, B33-B42.

McQueen, J. M. (in press). Speech perception. In K. Lamberts, & R. Goldstone (Eds.), *The handbook of cognition*. London: Sage Publications.

McQueen, J. M., Dahan, D. & Cutler, A. (in press). Continuity and gradedness in speech processing. In N. O. Schiller, & A. S. Meyer (Eds.), *Phonetic and Phonology in Language Comprehension and Production. Differences and Similarities* (pp. 37-76). Berlin, New York: Mouton de Gruyter.

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**, 1363-1389.

Mehler, J. (1981). The role of syllables in speech processing: Infant and adult data. *Philosophical Transactions of the Royal Society, Series B*, *295*, 333-352.

Milberg, W., Blumstein, S. & Dworetzky, B. (1988). Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, **26**, 305-308.

Miller, J. L. (1981). Effects of speaking rate on segmental distinctions. In P. D. Eimas, & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp. 39-74). Hillsdale, NJ: Erlbaum.

Miller, J. L. & Liberman, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception & Psychophysics*, **25**, 457-465

.

Miller, J. L. & Volaitis, L. E. (1989). Effects of speaking rate on the perceived internal structure of phonetic categories. *Perception & Psychophysics*, **46**, 505-512.

Mohr, B. (1971). Intrinsic variations in speech signal. *Phonetica*, **23**, 65-93.

Nearey, T. M. (1990). The segment as a unit of speech perception. *Journal of Phonetics*, **18**, 347- 373.

Nearey, T. M. (2001). Phoneme-like units and speech perception. *Language and Cognitive Processes*, **16**, 673-681.

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

Norris, D., Cutler, A., McQueen, J. M. & Butterfield, S. (submitted). Lexical activation in speech comprehension. Manuscript submitted for publication.

Norris, D., McQueen, J. M. & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, **23**, 299-325.

Norris, D., McQueen, J. M. & Cutler, A. (2002). Bias effects in facilitatory phonological priming. *Memory & Cognition*, **52**, 399-411.

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In *The production of speech* (P. F. MacNeilage, editor). New York: Springer-Verlag.

Ohde, R. N. (1984). Fundamental-frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America*, **75**, 224-230.

Perea, M., Rosa, E. & Gomez, C. (2002). Is the go/no-go lexical decision task an alternative to the yes/no lexical decision task? *Memory & Cognition*, **30**, 34-45.

Port, R. F. (1979). Influence of tempo on stop closure duration as a cue for voicing and place. *Journal of Phonetics*, **7**, 45-56.

Praamstra, P., Meyer, A. S. & Levelt, W. J. M. (1994). Neurophysiological manifestations of phonological processing: Latency variation of a negative ERP component timelocked to phonological mismatch. *Journal of Cognitive Neuroscience*, **6**, 204-219.

Radeau, M., Besson, M. Fonteneau, E. & Castro, S.L. (1998). Semantic, repetition and rime priming between spoken words: Behavioral and electrophysiological evidence. *Biological Psychology*, **48**, 183-205.

Radeau, M., Morais, J. & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 1297-1311.

Radeau, M., Segui, J. & Morais, J. (1994). The effect of overlap position in phonological priming between spoken words. *Proceedings of the 1994 International Conference on Spoken Language Processing*, **3**, 1419-1422. Yokohama: Acoustical Society of Japan.

Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, **92**, 81-110.

Repp, B. H. & Liberman, A. M. (1987). Phonetic category boundaries are flexible. In S. R. Harnad (Ed.), *Categorical perception* (pp. 89-112). Cambridge: Cambridge University Press.

Rothenberg, M. (1968). The breath-stream dynamics of simple-released-plosive production. *Bibliotheca Phonetica,* **6**.

Rothman, G. B., Koenig, L. L. & Lucero, J. C. (2002). Intraoral pressure trajectories during voiced and voiceless stops in women and children. *Journal of the Acoustical Society of America*, **112**, 2416.

Salverda, A. P., Dahan, D. & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, **90**, 51-89.

Slis, I. H. & Cohen, A. (1969). On complex regulating voiced-voiceless distinction I. *Language and Speech*, **12**, 80-102.

Slowiaczek, L. M. & Hamburger, M. B. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18**, 1239-1250.

Slowiaczek, L. M. & Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory & Cognition*, **14**, 230-237.

Slowiaczek, L.M., McQueen, J.M., Soltano, E.G. & Lynch, M. (2000). Phonological representations in prelexical speech processing: Evidence from form-based priming. *Journal of Memory and Language*, **43**, 530-560.

Smith, B. (1978). Effects of place of articulation and vowel environment on voiced stop consonant production. *Glossa*, **12**, 163-175.

Smits, R. (1995). *Detailed versus gross spectro-temporal cues for the perception of stop consonants*. Unpublished doctoral dissertation, Institute for Perception Research, Eindhoven.

Smits, R. (2000). Temporal distribution of information for human consonant recognition in VCV utterances. *Journal of Phonetics*, **28,** 111-135.

Smits, R. (2001). Evidence for hierarchical categorization of coarticulated phonemes. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 1145-1162.

Smits, R. L., Ten Bosch, L. & Collier, R. (1996). Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants II. Modeling and evaluation. *Journal of the Acoustical Society of America*, **100**, 3865-3881.

Smits, R., Warner, N., McQueen, J.M. & Cutler, A. (2003). Unfolding of phonetic information over time: A database of Dutch diphone perception. *Journal of the Acoustical Society of America*, **113**, 563-574.

Soloman, R. L. & Postman, L. (1952). Frequency of usage as a determinant of recognition thresholds for words. *Journal of Experimental Psychology,* **43**, 195-201.

Spinelli, E., McQueen, J. M. & Cutler, A. (2003). Processing resyllabified words in French. *Journal of Memory and Language*, **48**, 233-254.

Stevens, K. N. (1993). Models for the production and acoustics of stop consonants. *Speech Communication*, **13**, 367-375.

Stevens, K. N. (1998). *Acoustic Phonetics.* Cambridge: MIT Press.

Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America*, **111**, 1872-1891.

Stone, G. O. & Van Orden, G. C. (1993). Strategic control of processing in visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 744-774.

Streeter, L. A. & Nigro, G. N. (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America*, **65**, 1533-1541.

Studebaker, G. A. (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, **28**, 455-462.

Summerfield, A. Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1074-1095.

Svirsky, M. A., Stevens, K. N., Matthies, M.L., Manzella, J., Perkell, J. S. & Wilhelms-Tricarico, R. (1997). Tongue surface displacement during obstruent stop consonants. *Journal of the Acoustical Society of America* **102**, 562-571.

Tabossi, P., Collina, S., Mazzetti, M. & Zoppello, M. (2000). Syllables in the processing of spoken Italian. *Journal of Experimental Psychology: Human Perception and Performance*, **26**, 758-775.

Talkin, D. (1995). A Robust Algorithm for Pitch Tracking (RAPT). In Kleijn, W.B. & Paliwal, K.K. (Eds.), *Speech Coding and Synthesis*. New York: Elsevier.

Trueswell, J. C., Tanenhaus, M. K. & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **19**, 528-553.

Umeda, N. (1981). Influence of segmental factors on fundamental-frequency in fluent speech. *Journal of the Acoustical Society of America*, **70**, 350-355.

Utman, J. A. (1997). *Effects of subphonetic acoustic differences on lexical access in neurologically intact adults and patients with Broca's aphasia*. Unpublished doctoral dissertation, Brown University.

Utman, J. A., Blumstein, S. E. & Burton, M. W. (2000). Effects of subphonetic and syllable structure variation on word recognition. *Perception & Psychophysics*, **62**, 1297-1311.

Vitevitch, M.S. & Luce, P.A. (1999). Probabilistic phonotactics and spoken word recognition. *Journal of Memory & Language*, **40**, 374-408.

Westbury, J. R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America*, **73**, 1322-1336.

Westbury, J. R. & Keating, P. A. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, **22**, 145-166.

Whalen D. H., Abramson, A. S., Lisker, L. & Mody, M. (1993). F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustical Society of America*, **47**, 36-49.

Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics*, **35**, 49-64.

Whalen, D. H. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics*, **50**, 351-360.

Yeni-Komshian, G. H., Caramazza, A. & Preston, M. S. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics,* **5**, 35-48.

Yoshioka, H., Murase, S. & Uematsu, M. (1996). Palato-lingual contact patterns during voiced and unvoiced consonant production. *Journal of the Acoustical Society of America*, **100**, 2661.

APPENDICES

APPENDIX 2-A

Materials used in Experiment 2.1

	Vowel Context		Consonant Context	
	Labial	Alveolar	Labial	Alveolar
NW -Comp	baag	daaf	bleep	dreek
	beucht	darf	blog	drens
	bimp	deust	breld	dweum
	borf	dorg	brim	dwomp
NW +Comp	bark	daart	bluim	draan
	bech	dest	bluk	droost
	biek	dint	brijs	dwaalf
	bijn	doon	broef	dwijg
W -Comp	bars *(grim)*	dag *(day)*	bloem *(flower)*	draad *(thread)*
	beest *(beast)*	deur *(door)*	blos *(blush)*	draf *(trot)*
	berg *(mountain)*	dons *(down)*	bries *(breeze)*	dwaas *(foolish)*
	biels *(sleeper)*	duik *(dive)*	brood *(bread)*	dwars *(diagonal)*
W +Comp	baars *(perch)*	dak *(roof)*	blaag *(brat)*	drab *(dregs)*
	bed *(bed)*	dolk *(dagger)*	blad *(leaf)*	drek *(muck)*
	bink *(hunk)*	doorn *(thorn)*	bril *(glasses)*	drol *(turd)*
	boot *(boat)*	duin *(dune)*	brul *(roar)*	druk *(pressure)*

APPENDIX 2-B

Materials used in Experiment 2.2

	Vowel Context		Consonant Context	
	Labial	Alveolar	Labial	Alveolar
NW +Comp	bark-park *-(park)*	daart-taart *-(pie)*	bluim-pluim *-(feather)*	draan-traan *-(tear)*
	bech-pech *-(bad luck)*	dest-test *-(test)*	bluk-pluk *-(tuft)*	droost-troost *-(comfort)*
	biek-piek *-(peak)*	dint-tint *-(hue)*	brijs-prijs *-(price)*	dwaalf-twaalf *-(twelve)*
	bijn-pijn *-(pain)*	doon-toon *-(tone)*	broef-proef *-(trial)*	dwijg-twijg *-(twig)*
W -Comp	balk-palk *(beam)-*	damp-tamp *(vapour)-*	blok-plok *(block)-*	draad-traad *(thread)-*
	beek-peek *(brook)-*	deugd-teugt *(virtue)-*	brood-prood *(bread)-*	drop-trop *(liqourice)-*
	beurs-peurs *(grant)-*	dons-tons *(down)-*	braam-praam *(blackberry)-*	dwaas-twaas *(foolish)-*
	borg-porg *(bail)-*	duim-tuim *(thumb)-*	breed-preed *(broad)-*	dwerg-twerg *(dwarf)-*
W +Comp	baars-paars *(perch)-(purple)*	dak-tak *(roof)-(branch)*	blaag-plaag *(brat)-(plague)*	drab-trap *(dregs)-(stairs)*
	bek-pek *(mouth)-(tar)*	dolk-tolk *(dagger)-(interpreter)*	blad-plat *(leaf)-(flat)*	drek-trek *(muck)-(pull)*
	beul-peul *(brute)-(pod)*	doorn-toorn *(thorn)-(anger)*	blind-plint *(blind)-(skirting)*	drol-trol *(turd)-(troll)*
	bink-pink *(hunk)-(pinkie)*	duin-tuin *(dune)-(garden)*	bril-pril *(glasses)-(young)*	druk-truck *(pressure)-(truck)*

Note that some of the items which were used in the nonword conditions of Experiments 2.1 and 2.2 were very low frequent words which were expected to be unknown to the participants.

APPENDIX 3-A

Materials used in Experiments 3.1, 3.2, and 3.3

HIGH FREQUENCY PRIMES

Related PRIME	Unrelated PRIME	TARGET (Exp 3.1)
baby *(baby)*	basis *(basis)*	kind *(child)*
bijbel *(bible)*	bakker *(baker)*	geloof *(religion)*
bang *(afraid)*	boog *(arch)*	angst *(fear)*
beeld *(screen)*	blond *(blond)*	televisie *(television)*
beest *(animal)*	blauw *(blue)*	dier *(animal)*
berg *(mountain)*	braaf *(good)*	dal *(valley)*
bitter *(bitter)*	burger *(citizen)*	zoet *(sweet)*
blij *(happy)*	blok *(block)*	vrolijk *(jolly)*
bloed *(blood)*	boer *(farmer)*	rood *(red)*
bloem *(flower)*	baan *(job)*	roos *(rose)*
bloot *(naked)*	buurt *(neighborhood)*	naakt *(naked)*
bodem (*ground)*	beker *(mug)*	grond *(ground)*
bos *(woods)*	brief *(letter)*	boom *(tree)*
brand *(fire)*	borst *(chest)*	vuur *(fire)*
breed *(wide)*	bron *(source)*	smal *(narrow)*
broek *(trousers)*	bank *(cough)*	riem *(belt)*
broer *(brother)*	blik *(tin)*	zus (*sister)*
brood *(bread)*	boek *(book)*	boter *(butter)*
brug *(bridge)*	beurt *(turn)*	rivier *(river)*
bruin *(brown)*	baas *(boss)*	zwart *(black)*
dag *(day)*	doel *(goal)*	nacht *(night)*
datum *(date)*	dronken *(drunk)*	kalender *(calender)*
deksel *(lid)*	dringend *(urgent)*	pan *(pan)*
deur *(door)*	dom *(stupid)*	raam *(window)*
dicht *(closed)*	dank *(thanks)*	open *(open)*
dienst *(duty)*	ding *(thing)*	leger *(army)*
dochter *(daugther)*	duiven *(pigeons)*	zoon *(son)*
dokter *(doctor)*	dreiging *(threat)*	arts *(doctor)*
donker (*dark)*	dertig *(thirty)*	licht *(light)*
dood *(dead)*	droom *(dream)*	levend *(alive)*
dorp *(village)*	dwars *(crosswise)*	stad *(city)*
douche *(shower)*	duivel *(devil)*	bad *(bath)*
draad *(thread)*	dief *(thief)*	naald *(needle)*
drank *(booze)*	diep *(deep)*	bier *(beer)*
drie *(three)*	doek *(cloth)*	vier *(four)*
druif *(grape)*	daad *(act)*	wijn *(wine)*
druppel *(drop)*	dapper *(brave)*	water *(water)*
dubbel *(double)*	droevig *(sad)*	twee *(two)*
duim *(thumb)*	doos *(box)*	vinger *(finger)*
duizend (*thousand)*	dame *(lady)*	getal *(number)*

Materials used in Experiments 3.1, 3.2, and 3.3

LOW FREQUENCY PRIMES

Related PRIME	Unrelated PRIME	TARGET (Exp 3.1)
baken *(beacon)*	biceps *(biceps)*	zee *(sea)*
bamboe *(bamboo)*	babbel *(chat)*	hout *(wood)*
banjo *(banjo)*	bastaard *(bastard)*	gitaar *(guitar)*
bars *(grim)*	braam *(blackberry)*	nors *(surly)*
batik *(batik)*	bode *(messenger)*	stof *(fabric)*
beitel *(chisel)*	bede *(prayer)*	hamer *(hammer)*
beits *(stain)*	broche *(brooch)*	verf *(paint)*
beta *(beta)*	balsem *(balsam)*	alfa *(alfa)*
biels *(sleeper)*	bof *(mumps)*	spoor *(track)*
bingo *(bingo)*	bever *(beaver)*	spel *(game)*
bizon *(bison)*	bloesem *(blossom)*	buffel *(buffalo)*
blaam *(blame)*	blits *(trendy)*	schuld *(guilt)*
bochel *(hump)*	bengel (*scamp)*	rug *(back)*
blos *(blush)*	bonk *(lump)*	wang *(cheek)*
blubber *(mud)*	blanco *(blank)*	modder *(mud)*
bokking *(smoked herring)*	bunker *(bunker)*	vis *(fish)*
bonje *(row)*	berging *(storeroom)*	ruzie *(fight)*
braille *(braille)*	bistro *(bistro)*	blind *(blind)*
bries *(breeze)*	bef *(jabot)*	wind *(wind)*
buidel *(pouch)*	beugel *(brace)*	kangoeroe *(kangaroo)*
buil *(bump)*	brem *(broom)*	bult *(lump)*
buizerd *(buzzard)*	balie *(counter)*	vogel *(bird)*
bumper *(bumper)*	brouwse*l (brew)*	auto *(car)*
bundel *(collection)*	bivak *(bivouac)*	gedicht *(poem)*
daalder *(Dutch coin)*	diva *(diva)*	gulden *(guilder)*
dadel *(date)*	drastisch *(drastic)*	vijg *(fig)*
diesel *(diesel)*	dogma *(dogma)*	benzine *(petrol)*
dille *(dill)*	drachtig *(bearing)*	kruid *(herb)*
distel *(thistle)*	drachme *(drachme)*	stekels *(prickles)*
donor *(donor)*	demper *(muffler)*	orgaan *(organ)*
dons *(down)*	duw *(push)*	zacht *(soft)*
dooier *(yolk)*	duster *(dressing gown)*	ei *(egg)*
draf *(trot)*	dump *(dump)*	paard *(horse)*
drassig *(swampy)*	divan *(divan)*	moeras *(swamp)*
dreumes *(toddler)*	delta *(delta)*	kleuter *(infant)*
drum *(drum)*	disk *(disk)*	trommel *(drum)*
duf *(stuffy)*	drang *(urge)*	moe *(tired)*
duik *(dive)*	dreun *(blow)*	zwemmen *(swimming)*
dupe *(dupe)*	duo *(duo)*	slachtoffer *(victim)*
duplo *(duplo)*	dekking *(cover)*	lego (*lego)*

APPENDIX 4-A

Materials used in Experiments 4.1, 4.2, and 4.3

BLUE CONDITION

Voiced (word)	Voiceless (nonword)	Unrelated (word)
beek *(brook)*	peek	beurs *(grant)*
beest *(beast)*	peest	brug *(bridge)*
berg *(mountain)*	perg	blond *(blond)*
berm *(verge)*	perm	buit *(loot)*
* biels *(sleeper)*	piels	beits *(stain)*
* big *(piglet)*	pig	brons *(bronze)*
blauw *(blue)*	plauw	buurt *(neighborhood)*
bleek *(pale)*	pleek	brand *(fire)*
* blits *(trendy)*	plits	brom *(buzz)*
bloem *(flower)*	ploem	baan *(job)*
blos *(blush)*	plos	bars *(stern)*
boer *(farmer)*	poer	barst *(crack)*
* bok *(goat)*	pok	bruid *(bride)*
* bonk *(chunck)*	ponk	brink *(village green)*
boor *(drill)*	poor	blaar *(blister)*
braam *(blackberry)*	praam	boeg *(bow)*
brein *(brain)*	prein	boog *(arch)*
brief *(letter)*	prief	boom *(tree)*
* bries *(breeze)*	pries	blaas *(bladder)*
broer *(brother)*	proer	bos *(woods)*
brok *(lump)*	prok	bloot *(naked)*
bron *(source)*	pron	bloei *(blossom)*
brood *(bread)*	prood	balk *(beam)*
bros *(brittle)*	pros	bult *(bump)*
* buil *(bump)*	puil	brie *(brie)*
* buis *(tube)*	puis	bom *(bomb)*
burcht *(castle)*	purcht	broos *(frail)*
damp *(vapour)*	tamp	dief *(thief)*
darm *(intestine)*	tarm	dwaas *(foolish)*
deur *(door)*	teur	drie *(three)*
* dicht *(closed)*	ticht	drank *(beverage)*
dons *(down)*	tons	duik *(dive)*
dorp *(village)*	torp	dag *(day)*
draad *(thread)*	traad	doel *(goal)*
draak *(dragon)*	traak	douche *(shower)*
drum *(drum)*	trum	deeg *(dough)*
duif *(pigeon)*	tuif	doek *(cloth)*
dwars *(diagonal)*	twars	dood *(death)*
dweil *(cloth)*	tweil	draai *(turn)*
dwerg *(dwarf)*	twerg	doof *(deaf)*

Materials used in Experiments 4.1, 4.2, and 4.3

BEAR CONDITION

	Voiced (word)	Voiceless (word)	Unrelated (word)
*	baal *(bale)*	paal *(pile)*	bef *(jabot)*
	baard *(beard)*	paard *(horse)*	broek *(trousers)*
	baars *(perch)*	paars *(purple)*	bluf *(brag)*
	bad *(bath)*	pad *(path)*	blij *(happy)*
	bak *(tray)*	pak *(suit)*	blok *(block)*
	band *(tire)*	pand *(building)*	blik *(tin)*
	beer *(bear)*	peer *(pear)*	breuk *(fraction)*
	berk *(birch)*	perk *(flowerbed)*	brits *(plank bed)*
	beuk *(beech)*	peuk *(stump)*	bruut *(brute)*
	beul *(brute)*	peul *(pod)*	biecht *(confession)*
	bijl *(axe)*	pijl *(arrow)*	boon *(bean)*
*	bink *(hunk)*	pink *(pinkie)*	bres *(breach)*
*	blaag *(brat)*	plaag *(plague)*	broom *(bromine)*
	blad *(leave)*	plat *(flat)*	braaf *(good)*
	blank *(white)*	plank *(shelf)*	beurt *(turn)*
*	boef *(knave)*	poef *(pouf)*	bes *(berry)*
	boel *(a lot)*	poel *(puddle)*	bruin *(brown)*
	boord *(collar)*	poort *(gate)*	bloed *(blood)*
*	boos *(angry)*	poos *(while)*	breed *(broad)*
	boot *(boat)*	poot *(paw)*	beeld *(image)*
	bot *(bone)*	pot *(jar)*	bang *(afraid)*
	bouw *(construction)*	pauw *(peacock)*	bar *(bar)*
*	brij *(porridge)*	prei *(leek)*	blaam *(blame)*
	bril *(glasses)*	pril *(young)*	bank *(bench)*
*	brul *(roar)*	prul *(bauble)*	bout *(bolt)*
	bul *(bull)*	pul *(jug)*	branche *(line)*
	buur *(neighbor)*	puur *(pure)*	borst *(chest)*
	dak *(roof)*	tak *(branch)*	duim *(thumb)*
	dam *(dam)*	tam *(tame)*	drang *(impulse)*
	das *(tie)*	tas *(bag)*	dek *(cover)*
	dof *(dull)*	tof *(cool)*	drift *(passion)*
	dol *(mad)*	tol *(toll)*	druif *(grape)*
	dolk *(dagger)*	tolk *(interpreter)*	dijk *(dike)*
*	dooi *(thaw)*	tooi *(ornament)*	deuk *(dent)*
*	doorn *(thorn)*	toorn *(anger)*	dwang *(compulsion)*
	dop *(cap)*	top *(top)*	dreun *(roar)*
	dor *(dry)*	tor *(beetle)*	draf *(trot)*
	drol *(turd)*	trol *(troll)*	duw *(push)*
	duin *(dune)*	tuin *(garden)*	deugd *(virtue)*
*	duit *(cent)*	tuit *(spout)*	disk *(disc)*

Materials used in Experiments 4.1, 4.2, and 4.3

BLEM CONDITION

	Voiced (nonword)	Voiceless (nonword)	Unrelated (nonword)
	baag	paag	beim
	bans	pans	bluif
	baun	paun	bork
	bemp	pemp	blans
	beus	peus	bamp
	biens	piens	braap
	bimp	pimp	buif
*	blaaf	plaaf	broen
	blarp	plarp	beuf
	bleep	pleep	boof
	blem	plem	burf
	blimp	plimp	beeuw
	boens	poens	beeg
	boop	poop	baug
	borf	porf	bem
	braas	praas	beig
	bralm	pralm	bien
	braup	praup	bieft
	breeg	preeg	bilm
	breun	preun	baust
	brim	prim	blaap
	broeg	proeg	baaf
	brolf	prolf	baam
	broon	proon	buuft
*	bruif	pruif	blig
	buip	puip	blors
	buug	puug	blef
	daaf	taaf	dirp
	dars	tars	dift
	deig	teig	duust
	delm	telm	dreeg
	deust	teust	dorf
	diest	tiest	daam
*	dreek	treek	doeng
	drens	trens	diem
	droof	troof	dweik
*	drooi	trooi	daust
*	duip	tuip	draam
*	dwaag	twaag	doemp
	dwam	twam	dooks

Materials used in Experiments 4.1, 4.2, and 4.3

BRINCE CONDITION

	Voiced (nonword)	Voiceless (word)	Unrelated (nonword)
	baus	paus *(pope)*	bolm
	bech	pech *(bad luck)*	bift
	bees	pees *(tendon)*	briek
*	bels	pels *(fur)*	bruig
	bens	pens *(paunch)*	bramp
*	bers	pers *(press)*	bluig
	biek	piek *(peak)*	braalf
	bijn	pijn *(pain)*	brog
	bijp	pijp *(pipe)*	brong
	blaats	plaats *(place)*	barf
	blak	plak *(slice)*	beift
	blant	plant *(plant)*	boest
	blein	plein *(square)*	bruf
	blens	plens *(splash)*	birf
*	blons	plons *(spatter)*	brieg
	blooi	plooi *(fold)*	buint
	bluim	pluim *(feather)*	bieg
	bluis	pluis *(fluff)*	brimp
	bluk	pluk *(tuft)*	braft
	boes	poes *(cat)*	blauk
*	bols	pols *(wrist)*	brump
	bret	pret *(fun)*	balp
	briem	priem *(awl)*	blook
	brins	prins *(prince)*	baats
	broef	proef *(trial)*	bleug
	buin	puin *(rubbish)*	blerk
	bunt	punt *(point)*	bref
*	daai	taai *(tough)*	drelp
	dand	tand *(tooth)*	dwes
*	deil	teil *(pan)*	daft
	dekst	tekst *(text)*	dramp
	dest	test *(test)*	drak
	dint	tint *(hue)*	draast
*	dong	tong *(tongue)*	def
	doost	toost *(toast)*	drauf
*	drein	trein *(train)*	dulf
	droon	troon *(throne)*	deets
	drui	trui *(sweater)*	dosp
	dulp	tulp *(tulip)*	diecht
*	dwijg	twijg *(twig)*	deuf

Note: Items marked with an asterisk were not included in any analyses.

SAMENVATTING

PERCEPTUELE RELEVANTIE VAN PREVOICING IN HET NEDERLANDS

> ***La parole est moitié à celuy qui parle, moitié à celuy qui l'escoute;***
> ***cettuy cy se doibt preparer à la recevoir, selon le bransle qu'elle prend.***
>
> Spraak behoort voor de helft toe aan de spreker, voor de helft aan de luisteraar;
> de laatste moet zich erop voorbereiden het te ontvangen, in overeenstemming met de dynamiek die het heeft.
>
> MICHEL DE MONTAIGNE, LES ESSAIS III, PARIS (1588)

Natuurkundig gezien bestaat het spraaksignaal uit niets anders dan verdichtingen en verdunningen van de lucht. Als een persoon spreekt, verandert ze tijdelijk de druk van de lucht om haar heen. In deze luchtdrukveranderingen zit de boodschap van de spreker opgeslagen. Om de betekenis van deze boodschap te begrijpen moet de luisteraar weten hoe deze luchtdrukveranderingen moeten worden gedecodeerd. Dan pas krijgen de spraakklankpatronen betekenis. Aangezien het spraaksignaal van nature vergankelijk is (de luchtdrukverschillen verdwijnen immers snel) moet het decoderen snel plaatsvinden. Luisteraars zijn verbazingwekkend goed en snel in het analyseren en decoderen van het spraaksignaal. De luisteraar is in staat om in korte tijd de relevante informatie uit het spraaksignaal te halen en deze akoestische eigenschappen te vergelijken met linguïstische vormen die liggen opgeslagen in de hersenen en zo de bedoelde woorden te herkennen. Dit matchingsproces tussen het inkomende spraaksignaal en de opgeslagen lexicale representaties is niet zo eenvoudig als men zou verwachten op grond van het gemak waarmee luisteraars spraak begrijpen.

Een eigenschap van het spraaksignaal is dat het zeer variabel is. Neem bijvoorbeeld de spraakklank /s/. Deze wordt niet altijd op dezelfde manier uitgesproken. Verschillende sprekers produceren deze klank op een andere wijze, want allerlei factoren zoals het geslacht en leeftijd van de spreker, de vorm van het spraakkanaal, het dialect van de spreker, de spreeksnelheid en spreekstijl hebben invloed op de precieze akoestische realisatie van deze klank. Daarnaast zijn er nog allerlei sprekeronafhankelijke factoren, zoals de omgeving waarin de spreker zich bevindt. Bovendien hebben de omliggende klanken een relatief grote invloed op de fonetische realisatie van een spraakklank. Het foneem /s/ klinkt bijvoorbeeld duidelijk anders in een woord als *soep* dan in een woord als *sip*. Dit komt doordat de spreker zich tijdens de productie van de /s/ al voorbereidt op de volgende klinker. De articulatie van de /s/ wordt dus beïnvloed door de klinker /u:/ of /ɪ/. Tijdens de productie van de /s/ in *soep* zijn de lippen al gerond als voorbereiding op de /u:/, terwijl de lippen tijdens de /s/ in *sip* gespreid zijn. Dit verschijnsel noemen we coarticulatie. Coarticulatie vindt ook plaats over woordgrenzen, waardoor de /s/ in *soep* anders is in de zin *ik wil soep* dan in de zin *ik neem soep*. Zelfs als een woord meerdere keren door dezelfde spreker in dezelfde omgeving en context wordt uitgesproken, zal iedere uiting akoestisch verschillen.

Als gevolg van de variabiliteit van spraak en het feit dat er maar een beperkte fonetische ruimte is waarbinnen akoestische eigenschappen kunnen variëren, vertonen bepaalde spraakklanken behoorlijke fonetische overlap. Neem bijvoorbeeld de fonemen /b/ en /p/. Hoewel deze klanken fonemisch verschillen (het woord *baard* heeft immers een andere betekenis dan het woord *paard*), lijken ze fonetisch heel erg op elkaar. Een van de akoestische aanwijzingen (een akoestische cue) voor het verschil tussen /b/ en /p/ in het midden van een woord, zoals in de woorden *aanpakken* en *aanbakken*, is de duur van de sluiting. Als je een plofklank zoals /b/ of /p/ uitspreekt sluit je eerst alle luchtopeningen naar buiten af, zodat er geen lucht kan ontsnappen. Als gevolg van deze sluiting neemt de druk in de mondholte razendsnel toe. Vervolgens wordt de sluiting opgeheven, hetgeen resulteert in een ruisplof. De duur van de sluiting is over het algemeen langer voor stemloze plofklanken (zoals /p/) dan voor stemhebbende plofklanken (zoals /b/). De duur van deze sluiting varieert echter met het spreektempo. Daardoor kan het voorkomen dat de het verschil in sluitingsduur groter is tussen een langzaam en snel uitgesproken versie van het woord *aanpakken* dan tussen de woorden *aanpakken* en *aanbakken*. De luisteraar moet weten dat het eerste duurverschil geen foneemverschil aangeeft (in beide gevallen bedoelde de spreker *aanpakken*), terwijl het tweede, kleinere, duurverschil wel een fonemisch verschil aangeeft. Luisteraars moeten dus weten welk akoestisch detail relevant is voor het onderscheid tussen fonemen, en dus tussen woorden, en welk detail niet relevant is. Dit proefschrift gaat grotendeels over de manier waarop het menselijk spraakherkenningsysteem omgaat met relevant en irrelevant akoestische detail tijdens het proces van woordherkenning.

De akoestische eigenschap die ik in dit proefschrift onderzoek is *prevoicing*. Prevoicing is de aanwezigheid van stembandtrilling tijdens de sluiting van een plofklank aan het begin van een woord. In het Nederlands worden de stemhebbende plofklanken /b/ en /d/ geproduceerd met prevoicing, terwijl de stemloze plofklanken /p/ en /t/ zonder prevoing worden uitgesproken. Om te begrijpen wat de perceptuele relevantie van prevoicing is, is het noodzakelijk om eerst te onderzoeken hoe prevoicing varieert in het Nederlands. Luisteraars blijken namelijk heel gevoelig te zijn voor de manier waarop spraakklanken variëren. Deze statistische informatie helpt de luisteraar bij de manier waarop de verschillende soorten akoestisch detail moeten worden gewogen tijdens het herkenningsproces. Als luisteraar kan je het best vertrouwen op akoestische eigenschappen die systematisch voorkomen in bepaalde spraakklanken.

Het doel van Hoofdstuk 2 was om te bepalen hoe prevoicing varieert in het Nederlands. In het eerste experiment van dit hoofdstuk werden 10 sprekers gevraagd om 64 items uit te spreken, die begonnen met een /b/ of een /p/. Uit de resultaten bleek dat maar liefst 25% van deze uitingen waren geproduceerd zonder prevoicing. Daarnaast bleek dat de duur van de prevoicing, indien aanwezig, behoorlijk varieerde. Verschillende factoren bleken invloed te hebben op het percentage uitingen dat met prevoicing was uitgesproken en op de duur daarvan. Mannelijke sprekers produceerden vaker prevoicing dan vrouwelijke sprekers. Dit verschil wordt waarschijnlijk veroorzaakt door het verschil in de grootte van het spraakkanaal. Mannen hebben over het algemeen een groter spraakkanaal dan vrouwen, waardoor het makkelijker is om de stembanden te laten trillen tijdens de sluiting van de plofklank. Het verschil in duur van de aanwezige prevoicing tussen mannen en vrouwen was echter niet significant. Verder waren er twee segmentele factoren die de aanwezigheid van prevoicing beïnvloedden, namelijk de plaats van articulatie van de plofklank en het type foneem dat op de plofklank volgde. Labiale plofklanken (/b/) werden vaker met prevoicing geproduceerd dan alveolaire plofklanken (/d/) en plofklanken gevolgd door een klinker hadden vaker prevoicing dan plofklanken gevolgd door een medeklinker. Deze laatste factor bleek ook een effect te hebben op de duur van de aanwezige prevoicing (langere prevoicing voor /b/ dan voor /d/). Deze segmentele effecten kunnen worden verklaard in termen van de mate waarin het spraaksignaal kan worden vergroot tijdens de sluiting van de plofklank. Het vergroten van het spraakkanaal, actief of passief, maakt het produceren van prevoicing makkelijker omdat daardoor de druk boven de glottis minder snel stijgt. De sluiting van de plofklank /b/ is verder naar voren in de mond dan de sluiting van de plofklank /d/. Daardoor is er een groter oppervlak vrij dat kan deelnemen aan de passieve expansie. In het geval van de alveolaire plofklanken, kunnen de wanden van de farynx en een deel van het zachte verhemelte naar buiten worden geduwd als reactie op de stijgende druk, terwijl bij de productie van labiale plofklanken deze oppervlakken en delen van de wangen kunnen meedoen aan de expansie. Het effect van de aangrenzende klank komt hoogstwaarschijnlijk ook door het verschil in vrijheid om het spraakkanaal te vergroten. Wanneer een plofklank wordt gevolgd door een klinker is er waarschijnlijk meer mogelijkheid tot expansie dan wanneer de plofklank wordt gevolgd door een medeklinker. De exacte mechanismen van expansie in deze verschillende contexten zijn echter minder duidelijk. De twee lexicale factoren die waren onderzocht, de lexicale status van het item (*berg* is een woord; *beucht* is een nonsenswoord) en de aanwezigheid van een stemloze woordconcurrent (*berg* heeft geen concurrent, want *perg* bestaat niet; *baars* heeft wel een concurrent, want *paars* is ook een woord) hadden geen effect op de aanwezigheid of duur van prevoicing.

Het verschil tussen fonemen, en dus tussen woorden, wordt door meerdere akoestische cues wordt aangeduid. Sommige van deze cues zijn belangrijker zijn dan andere. De bevinding dat een kwart van de stemhebbende plofklanken werd geproduceerd zonder prevoicing doet de vraag rijzen hoe deze uitingen worden waargenomen door luisteraars. De afwezigheid van prevoicing zou er immers op duiden dat het foneem stemloos is. Het kan echter ook zijn dat prevoicing niet zo’n sterke cue is als algemeen wordt aangenomen. In het tweede experiment dat wordt beschreven in Hoofdstuk 2, werd onderzocht welke andere akoestische eigenschappen, naast prevoicing, het signaal heeft en welke van deze eigenschappen het best zouden kunnen worden gebruikt om het verschil tussen stemhebbende en stemloze plofklanken waar te nemen. Hiervoor werden 48 stemhebbend-stemloos paren gebruikt, die door 10 verschillende sprekers waren geproduceerd (in totaal 960 uitingen). Voor al deze uitingen werden de volgende akoestische maten verkregen: duur van de prevoicing, duur van de plof, het vermogen van de plof boven 500 Hz, het spectrale zwaartepunt van de plof, de F0 onmiddellijk na de plof en het verschil tussen de F0 onmiddellijk na de plof en de F0 in het midden van de (medeklinker plus) klinker. Al deze maten, behalve de F0 na de plof, toonden een significant verschil tussen stemhebbende en stemloze plofklanken. Een CART-analyse liet zien dat de af- of aanwezigheid van prevoicing de beste voorspeller is voor het verschil tussen stemhebbende en stemloze plofklanken. Om er achter te komen of luisteraars inderdaad het meest vertrouwen op de af- of aanwezigheid van prevoicing tijdens de waarneming van deze klanken werd er een derde experiment uitgevoerd. Hierbij kregen 16 luisteraars alle 960 uitingen te horen en moesten beslissen of het item begon met een /b/ of een /p/, of met een /d/ of een /t/. Een tweede CART-analyse, dit maal op de proportie stemhebbende responsen van iedere uiting, liet zien dat luisteraars inderdaad het meest afgingen op de af- of aanwezigheid van prevoicing. Alle plofklanken die waren geproduceerd met prevoicing werden als stemhebbend waargenomen. Het grootste deel van de uitingen die waren geproduceerd zonder prevoicing werden als stemloos waargenomen (in deze groep bevinden zich zowel de bedoelde stemloze plofklanken en de bedoelde stemhebbende plofklanken zonder prevoicing). Een deel van de uitingen zonder prevoicing werd echter als stemhebbend waargenomen, op grond van de andere akoestische eigenschappen. Dit laat zien dat een /b/ of een /d/ zonder prevoicing nog steeds als zodanig kan worden herkend. De invloed van deze andere akoestische eigenschappen verschilde voor de twee plaatsen van articulatie: voor de labiale plofklanken zonder prevoicing bleek het F0 verschil de belangrijkste cues te zijn, terwijl het spectrale zwaartepunt de meeste invloed had bij de alveolaire plofklanken zonder prevoicing. De invloed van deze secundaire cues is echter zwak in vergelijking met de invloed van de primaire cue prevoicing.

De bevinding dat prevoicing de sterkste perceptuele cue is voor stemhebbende plofklanken, hoewel prevoicing regelmatig afwezig is in de productie van stemhebbende plofklanken, maakt de vraag naar de effecten van variatie van prevoicing op lexicale activatie nog interessanter. Aan de ene kant zou men kunnen voorspellen dat het weghalen van prevoicing een groot effect heeft op de activatie van woorden die met dit foneem beginnen. Door de prevoicing weg te halen, haal je de belangrijkste cue voor stemhebbendheid weg. De meerderheid van de plofklanken zonder prevoicing die je hoort zijn immers stemloos bedoeld. Aan de andere kant blijkt dat stemhebbende plofklanken regelmatig zonder prevoicing worden uitgesproken. Luisteraars zijn in hun leven dus regelmatig woorden tegengekomen die met een stemhebbende plofklank beginnen, maar zonder prevoicing zijn uitgesproken. Luisteraars hebben dus door ervaring geleerd dat prevoicing niet altijd aanwezig is. Het zou daarom kunnen dat de afwezigheid van prevoicing geen groot effect heeft op lexicale activatie.

In Hoofdstuk 3 en 4 werden de effecten van prevoicingvariatie op woordherkenning onderzocht. Er werden twee soorten variatie tegen over elkaar gesteld: een verschil in de aan- en afwezigheid van prevoicing (6 versus 0 perioden prevoicing) en een verschil in de exacte hoeveelheid prevoicing (12 versus 6 perioden prevoicing). Beide soorten verschillen in prevoicing zijn even groot; in beide gevallen gaat het om een verschil van 6 perioden. Bovendien vallen beide soorten binnen het normale bereik van prevoicingvariatie. De twee typen variatie verschillen echter in perceptuele relevantie. Zoals in Hoofdstuk 2 van dit proefschrift uitvoerig wordt beschreven, is de primaire cue voor het stemonderscheid in Nederlandse initiële plofklanken de af- of aanwezigheid van prevoicing. De exacte hoeveelheid prevoicing lijkt er niet toe te doen, zolang er maar prevoicing aanwezig is. De predictie was daarom dat het verschil tussen 6 en 0 perioden zou resulteren in een verschil in lexicale activatie van de woordkandidaten die met deze plofklanken beginnen, terwijl het verschil tussen 12 en 6 perioden geen verschil zou opleveren.

In hoofdstuk 3 worden de resultaten van twee primingexperimenten besproken, waarin het effect van de twee soorten prevoicingvariatie werden onderzocht. Het eerste experiment was een *cross-modal associative priming experiment*. In een dergelijk experiment krijgen proefpersonen eerst een woord of nonsenswoord te horen (de *prime*) en vervolgens een ander woord of nonsenswoord op het scherm te zien (de *target*). De taak van de proefpersoon is om zo snel mogelijk te beslissen of de target een woord is of niet, door zo snel mogelijk op de ja- of nee-knop te drukken. Als de prime qua betekenis gerelateerd is aan de target (bijvoorbeeld *bloem-roos*) zijn proefpersonen sneller in het maken van een lexicale decisie dan wanneer de prime en de target niet gerelateerd zijn (bijvoorbeeld *baan-roos).* Dit komt doordat bij het horen van het woord *roos* ook semantische gerelateerde woordkandidaten worden geactiveerd, waaronder *bloem*. Het verschil in snelheid tussen de twee condities, de gerelateerde en ongerelateerde, is dus een maat voor hoeveelheid activatie van het targetwoord als gevolg van het horen van de prime. De primes in het eerste primingexperiment waren of hoogfrequente woorden (zoals *bloem*) of laagfrequente woorden (zoals *biels*), die begonnen met een stemhebbende plofklank. De primes waren semantisch gerelateerd aan de target (*bloem-roos*) of ongerelateerd (*baan-roos*). Iedere gerelateerde prime kwam met drie verschillende hoeveelheden prevoicing voor: met 12, 6 of 0 perioden prevoicing. De resultaten lieten een keurig primingeffect zien: de lexicale decisies waren sneller wanneer de target was voorafgegaan door een gerelateerde prime dan wanneer de target was voorafgegaan door een ongerelateerde prime. Er was echter geen verschil tussen de drie prevoicingcondities. Het maakte niet uit of de prime met 12, 6 of 0 perioden prevoicing begon, in alle drie de gevallen waren proefpersonen even snel met beslissen dat de gerelateerde target een woord was. Dit suggereert dat de drie prevoicingvariaties resulteren in een zelfde hoeveelheid lexicale activatie. Dezelfde materialen werden vervolgens gebruikt in een tweede primingexperiment, dit maal een *cross-modal identity priming experiment*. Naast de hoog- en laagfrequente primes werden er nu ook nonsenswoord primes gebruikt. In dit experiment zijn de prime en de target identiek (*bloem-bloem*) of niet (*baan-bloem*). De resultaten lieten wederom een duidelijk primingeffect zien voor de identieke primes, maar er was weer geen effect van prevoicingvariatie. Het derde experiment, een discriminatie-experiment, liet zien dat luisteraars wel degelijk het verschil tussen de drie prevoicingvariaties konden horen. Het uitblijven van een effect in de primingtaken werd dus niet veroorzaakt door een onvermogen van de luisteraars om de verschillen waar te nemen.

De bevinding dat primes met initiële stemhebbende plofklanken zonder prevoicing resulteerden in een vergelijkbare mate van lexicale activatie als de plofklanken met prevoicing zou het resultaat kunnen zijn van het feit dat prevoicing in het Nederlands regelmatig afwezig is. Dus alhoewel luisteraars sterk vertrouwen op de aanwezigheid van prevoicing, hebben ze ook geleerd dat de afwezigheid van prevoicing niet direct betekent dat de plofklank stemloos is. Terwijl de aanwezigheid van prevoicing zonder twijfel aangeeft dat de plofklank stemhebbend is, geeft de afwezigheid van prevoicing aan dat de plofklank stemhebbend of stemloos is. Als dit inderdaad het geval is, zou men voorspellen dat plofklanken zonder prevoicing zowel de stemhebbende als de stemloze prelexicale representatie (bijvoorbeeld zowel /b/ als /p/) activeren en dat zowel lexicale kandidaten beginnend met een stemhebbende plofklank als lexicale kandidaten beginnend met een stemloze plofklank worden geactiveerd. Het verwijderen van de aanwezige prevoicing in stemhebbende plofklanken resulteert echter niet in een compleet ambigue klank. Alle andere akoestische cues zijn immers nog steeds aanwezig en geven aan dat het om een stemhebbende plofklank gaat. Dit werd bevestigd door de resultaten van de identificatietaken die volgden op beide primingexperimenten. Het percentage stemhebbende responsen was lager voor de stemhebbende plofklanken zonder prevoicing, maar niettemin werd het merendeel van deze uitingen nog steeds als stemhebbend waargenomen. De stemhebbende plofklanken zonder prevoicing zijn dus nog steeds stemhebbend, maar minder stemhebbend dan de stemhebbende plofklanken met prevoicing. De verwachting is dus dat de uitingen zonder prevoicing zowel de stemhebbende als de stemloze representaties zouden activeren, maar de eerste sterker dan de laatste. Een belangrijke eigenschap van de items in de primingexperimenten van Hoofdstuk 3 is dat geen van deze items een stemloze lexicale concurrent had. Bij het horen van bijvoorbeeld *bloem* zonder prevoicing, was *bloem* de enige serieuze lexicale kandidaat, aangezien *ploem* niet bestaat.

In Hoofdstuk 4 werd de invloed van de aanwezigheid van stemloze lexicale concurrenten onderzocht in twee cross-modal identity primingexperimenten. De primes waren woorden of nonsenswoorden met een initiële stemhebbende plofklank, waarvoor wel of geen stemloze lexicale concurrent bestond. Dit resulteerde in 4 condities, die gelabeld waren als de *Blauw-conditie* (woord zonder concurrent), de *Beer-conditie* (woord met concurrent), de *Blem-conditie* (nonsenswoord zonder concurrent) en de *Brins-conditie* (nonsenswoord zonder concurrent). Net als in de vorige primingexperimenten waren er drie prevoicing-condities (12, 6 en 0 perioden) en een ongerelateerde primingconditie. Nu was er echter ook een stemloze primingconditie, waarin de stemloze tegenhangers van de stemhebbende primes werden gepresenteerd (bijvoorbeeld *ploem*). In het eerste primingexperiment fungeerden de stemhebbende items als target (bijvoorbeeld *blauw*, *beer*, *blem* en *brins*) en in het tweede experiment dienden de stemloze items als target (bijvoorbeeld *plauw*, *peer*, *plem* en *prins*). Uit de resultaten bleek dat de lexicale decisies in alle condities sneller waren wanneer de prime en de target identiek waren. Er was geen facilitatie wanneer de prime en de target enkel qua eerste foneem verschilden, bijvoorbeeld *beer* en *peer* (of andersom).

De interessante bevinding was dat er nu een effect van prevoicingdeletie was in de condities met een stemloze lexicale concurrent (de Beer- en Brins-conditie). In de Beer-conditie, bijvoorbeeld, waren proefpersonen sneller om te beslissen dat de target *peer* een woord is wanneer ze net *beer* zonder prevoicing hadden gehoord, dan wanneer ze net *beer* met prevoicing hadden gehoord. Ze waren echter nog sneller wanneer ze net peer hadden gehoord. In de Brins-conditie, waren proefpersonen langzamer in het beslissen dat *brins* geen woord is wanneer ze net *brins* zonder prevoicing hadden gehoord, dan wanneer ze net *brins* met prevoicing hadden gehoord. Deze resultaten suggereren dat de primes zonder prevoicing zowel de stemloze als de stemhebbende woordkandidaten hadden geactiveerd. De laatste werd echter zwakker geactiveerd dan de eerste, omdat een stemhebbende plofklank zonder prevoicing nog altijd meer stemhebbend dan stemloos is. Dus het weghalen van prevoicing resulteert in graduele lexicale activatie van de stemloze lexicale concurrent. Deze graduele activatie was echter alleen maar meetbaar wanneer en conflict was tussen het visuele bewijs geleverd door de target en het sterkst door de prime geactiveerde woord. De graduele activatie van *peer* als gevolg van de presentatie van *beer* zonder prevoicing kon dus alleen worden gemeten wanneer de target *peer* was. In het experiment waar de target *beer* was werd geen graduele activatie gevonden.

De belangrijkste bevinding is echter dat het verschil tussen 6 en 0 perioden prevoicing resulteert in verschillen in lexicale activatiepatronen, terwijl er in geen van de condities een verschil werd gevonden tussen 12 en 6 perioden prevoicing. Deze resultaten werden versterkt door de resultaten van het tweede discriminatie-experiment. Luisteraars kregen in dit experiment ieder item uit alle 6 prevoicingcombinaties te horen. Ze moesten beslissen of de twee items precies hetzelfde klonken of net iets anders. Net als uit het eerste experiment bleek dat luisteraars het verschil tussen 12 en 0 perioden het beste konden horen, dat ze wat meer moeite hadden met het verschil tussen 6 en 0 perioden en dat ze het verschil tussen 12 en 6 perioden het moeilijkst vonden. Daarnaast werd er een duidelijk effect van aanwezigheid van een stemloze lexicale concurrent: het was makkelijker om het verschil te horen tussen twee verschillende hoeveelheden prevoicing als het item een stemloze woordconcurrent had, maar alleen als een van de twee leden van het paar geen prevoicing had. Dit suggereert opnieuw dat items met een stemhebbende plofklank zonder prevoicing ook de stemloze lexicale kandidaat significant activeren.

De bevindingen, die worden beschreven in dit proefschrift laten zien dat niet alle akoestische variatie invloed heeft op lexicale activatie. Luisteraars weten welk akoestisch verschil relevant is voor de herkenning van fonemen en van woorden en welk niet. Alleen de relevante informatie kan de activatie van lexicale kandidaten beïnvloeden. De relevantie van de verschillende akoestische eigenschappen hangt nauw samen met de manier waarop deze eigenschappen variëren in natuurlijke spraak. Luisteraars blijken zeer gevoelig te zijn voor dit soort distributionele informatie. Daarnaast laten de data een intrigerende paradox zien: prevoicing is de primaire cue voor het stemonderscheid in het Nederlands, maar tegelijkertijd is prevoicing regelmatig afwezig in de productie van stemhebbende plofklanken. Deze paradox wordt weerspiegeld in de manier waarop het spraakherkenningsysteem omgaat met stemhebbende plofklanken zonder prevoicing. Als er geen prevoicing is, wordt niet alleen de lexicale kandidaat beginnend met een stemhebbende plofklank geactiveerd, maar ook de kandidaat beginnend met een stemloze plofklank (als die er tenminste is). Niettemin belemmert de afwezigheid van prevoicing de luisteraar niet om het woord met een stemhebbende plofklank te herkennen.

De manier waarop Michel de Montaigne meer dan vier eeuwen geleden het fenomeen spraak beschrijft is dus zeer bekwaam. *La parole est moitié à celuy qui parle, moitié à celuy qui l'escoute...* Spraak behoort inderdaad zowel toe aan de spreker als de luisteraar. *Cettuy cy se doibt preparer à la recevoir, selon le bransle qu'elle prend*. De manier waarop de luisteraar zich moet voorbereiden om de boodschap van de spreker te begrijpen is door haar aandacht te vestigen op de dynamiek die spraak maakt. Op deze manier leert de luisteraar welk deel van de bewegingen zinvol is, welk deel niet zinvol is en welke bewegingspatronen vaker voorkomen dan andere. Deze informatie helpt de luisteraar de betekenis uit het complexe signaal te halen, dat door de spreker is geproduceerd.

CURRICULUM VITAE

Petra van Alphen werd op 9 februari 1975 geboren in Almelo. Na het behalen van haar vwo-diploma (gymnasiumstroom) aan de Openbare Scholengemeenschap Erasmus te Almelo besloot zij Algemene Letteren te gaan studeren aan de Universiteit Utrecht. In 1994 behaalde ze haar propedeuse en stapte ze over op de bovenbouwstudie Fonetiek. Tijdens haar studie gaf een Erasmusbeurs haar de mogelijkheid gedurende een semester Speech Communication te studeren aan de University of Sheffield. In oktober 1997 begon zij aan haar afstudeerproject in de Comprehension Group van het Max Planck Instituut voor Psycholinguïstiek, waar zij vanaf mei 1998 ook werkzaam was als onderzoeksassistent. In september 1998 behaalde ze haar doctoraaldiploma in de Fonetiek en bleef nog een paar maanden onderzoeksassistent op het MPI. In 1999 werd haar een stipendium toegekend door het Max-Planck-Gesellschaft om promotieonderzoek te doen binnen de Comprehension Group van het MPI. In 2001 kreeg zij, mede dankzij een reisbeurs van NWO, de kans om 5 maanden onderzoek te doen in het Speech Perception Lab van Northeastern University in Boston. Vanaf november 2002 maakt ze als junioronderzoeker deel uit van het project *Early language development in Specific Language Impairment and dyslexia: A prospective and comparative study* van de Universiteit Utrecht.

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 midsaggital articulography
   Niels O. Schiller

3. Lexical access in the production of ellipsis and pronouns
   Bernadette M. 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 Oceanic 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 Schultze-Berndt

15. Interpreting indefinites: 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: The selection of linking elements in Dutch compounds
Andrea Krott

19. Morphology in speech comprehension
Kerstin Mauth

20. Morphological families in the mental lexicon
Nivja H. de Jong

21. Fixed expressions and the production of idioms
Simone A. Sprenger

22. The grammatical coding of postural semantics in Goemai
Birgit Hellwig

23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies
Fermín Moscoso del Prado Martín

24. Contextual influences on spoken-word processing: An electrophysiological approach
Daniëlle van den Brink

25. Perceptual relevance of prevoicing in Dutch
Petra M. van Alphen