

Intraparadigmatic Effects on the Perception of voice

Mirjam Ernestus^{a,b} & R. Harald Baayen^{b,a}
^a *Max-Planck Institute for Psycholinguistics*
^b *Radboud University Nijmegen*

In Dutch, all morpheme-final obstruents are voiceless in word-final position. As a consequence, the distinction between obstruents that are voiced before vowel-initial suffixes and those that are always voiceless is neutralized. This study adds to the existing evidence that the neutralization is incomplete: neutralized, alternating plosives tend to have shorter bursts than non-alternating plosives. Furthermore, in a rating study, listeners scored the alternating plosives as more voiced than the non-alternating plosives, showing sensitivity to the subtle subphonemic cues in the acoustic signal. Importantly, the participants who were presented with the complete words, instead of just the final rhymes, scored the voiceless alternating plosives as even more voiced. This shows that listeners' perception of voice is affected by their knowledge of the obstruent's realization in the word's morphological paradigm. Apparently, subphonemic paradigmatic levelling is a characteristic of both production and perception. We explain the effects within an analogy-based approach.

1. Introduction

In many languages, morpheme-final obstruents alternate between voiced and voiceless, depending on their position in the word. Such obstruents are generally voiced before vowel-initial suffixes, and voiceless elsewhere (unless they are subject to regressive voice assimilation). To give an example, the final obstruent of the Dutch morpheme *mand* 'basket' is voiced in the plural [mɑndən] *manden* and voiceless in the singular [mant] *mand* (in this article we only give broad transcriptions). A word's morphological paradigm, which we define as consisting of all other words containing the word's stem, may thus show alternation in the voice specification of the stem-final obstruent.

This alternation of voice within morphological paradigms raises the question of whether the voiced stem-final obstruents affect the acoustic characteristics and interpretation of their intraparadigmatic voiceless counterparts. In other words, does the presence of [mɑndən] *manden* in the morphological paradigm of *mand* affect the production and interpretation of the [t] of [mant] *mand*?

In generative grammar, the alternation between voiced and voiceless obstruents is traditionally accounted for by means of underlying forms. A morpheme-final obstruent that is voiced before vowel-initial suffixes is also voiced in the underlying form. Thus, the underlying form of [mant] is /mand/ because of the voiced [d] in the plural [mɑndən]. Obstruents that are voiced in the underlying form are devoiced in syllable-final position by a rule or constraint of Final Devoicing (e.g., Booij 1995), which turns the /d/ of the singular /mand/ into [t]. Within generative grammar, the question of the effect of voice alternation on production and perception can thus be considered as a question of the effect of the underlying voice specification.

In this paper, we will not refer to the notion of underlying form. Recent studies have shown that the mental lexicon contains form representations for a great many words, including inflected forms (e.g., Jackendoff 1975, Baayen, Dijkstra & Schreuder 1997, Alegre & Gordon 1999, Baayen, McQueen, Dijkstra & Schreuder 2003). Thus, it contains the singular [mant] as well as the plural [mɑndən]. Given the storage of plurals, the realization of the stem-final obstruent in the plural need not be stored as a diacritic on the stem. It is already stored in the representation of the plural itself. We tentatively assume that all forms that are stored in the mental lexicon are surface representations, which reflect the actual realization of the forms, and that the lexicon does not contain abstract underlying forms. We refer to stem-final obstruents that surface as voiced in some forms of their paradigm as “alternating obstruents” (instead of “underlyingly voiced” obstruents), and we regard possible effects of voice alternation as intraparadigmatic effects, which the different forms of the same morphological paradigm may exert upon each other.

Previous studies have documented an effect of intraparadigmatic voice alternation in production for several languages, including Dutch, German, Polish, and Catalan (e.g., Dinnsen & Charles-Luce 1984, Port & O’Dell 1985, Slowiaczek & Dinnsen 1985, Port & Crawford 1989, Charles-Luce 1993, Warner, Jongman, Sereno & Kemps 2004, Ernestus & Baayen 2006). Although all word-final obstruents are generally voiceless in these languages, the alternating obstruents tend to have more acoustic characteristics of voiced obstruents than non-alternating obstruents, which are always voiceless. These alternating obstruents tend to be shorter, to be realized with vocal fold vibration during a longer period, and to be preceded by longer vowels. Thus, the [t] of a word such as [mant] *mand* ‘basket’ tends to have more acoustic characteristics of voiced obstruents than the [t] of a word such as [krant] *krant* ‘newspaper’, which has the plural [krantən] *kranten*. Hence, the neutralization at word-final position between alternating and non-alternating obstruents is incomplete. In what follows, we will refer to voiceless obstruents that possess some acoustic characteristics of genuine voiced obstruents, such as a relatively short duration or a relatively long preceding vowel, as weakly voiced.

For Dutch, the realization of alternating and non-alternating obstruents in existing words has been investigated by Warner et al. (2004). They carried out a production experiment with 15 native speakers of Dutch, and found a significant

difference in vowel duration of 3.5 ms. Ernestus & Baayen (2006) carried out a production experiment with pseudowords. The alternating/non-alternating character of final plosives was indicated by their spelling: In Dutch alternating plosives are represented by graphemes for voiced phonemes ([b] or [d]) in all positions in the word, while non-alternating plosives are represented by graphemes for voiceless phonemes ([p] or [t]). Ernestus and Baayen found a statistically significant difference between the alternating and non-alternating final plosives with respect to the duration of their release noise (their burst and the following period of aspiration). The release noise of plosives represented as voiceless were on average 23 ms longer than the release noises of plosives represented as voiced.

Several authors, however, argue that incomplete neutralization is not a characteristic of spontaneous speech, but would be induced by spelling, especially when speakers are asked to read aloud minimal word pairs such as *Rat–Rad* (e.g., Fourakis & Iverson 1984, Warner, Good, Jongman & Sereno 2006). Thus, Fourakis & Iverson (1984) reported a series of experiments in which speakers showed incomplete neutralization when reading aloud minimal word pairs, but failed to do so when they repeated infinitives and produced the corresponding past-tense forms and past participles. In contrast, Ernestus & Baayen (2006) revealed that speakers may also show incomplete neutralization when they read aloud lists of pseudowords without detecting the minimal word pairs and without being aware of the purpose of the experiment. Moreover, Dinnsen & Charles-Luce (1984) showed that incomplete neutralization is also present in Catalan minimal word pairs that do not reflect the difference between alternating and non-alternating obstruents in their spelling.

What is important for the present study is not so much whether speakers produce incomplete neutralization also in spontaneous conversations, but that listeners are sensitive to the fine acoustic differences induced by incomplete neutralization. Although the acoustic differences between alternating and non-alternating obstruents are generally small, listeners are able to take advantage of these subtle differences. They assign the correct spelling at significantly above chance level to the members of minimal word pairs that differ from each other only in the alternating/non-alternating character of the final obstruent (e.g., Port & O'Dell 1985, Port & Crawford 1989, Warner et al. 2004). For instance, when Dutch listeners hear [rat], they assign, at just above chance level, the intended meaning *raad* 'advice' with the plural [radən], or *raat* 'comb' with the plural [ratən]. They opt slightly more often for *raad* when the final obstruent is weakly voiced, and for *raat* when the obstruent is completely voiceless.

Additional evidence for listeners' sensitivity to incomplete neutralization comes from listeners' choices of allomorphs for pseudowords. In Dutch, the choice between the past tense allomorphs *–de* [də] and *–te* [tə] depends on the alternating/non-alternating character of the stem-final segment. If the segment is always realized as voiceless, the appropriate allomorph is *–te*, otherwise it is *–de* (see also Zonneveld, this volume). Ernestus & Baayen (2003) showed that when speakers do not know how the final obstruent is realized in morphologically related words, that is, when they have no information about whether it is an alternat-

ing obstruent, they tend to base their choice between *-te* and *-de* on the word's phonological similarity neighbourhood. If most words ending in the same type of rime take *-te*, speakers tend to choose *-te*, and if most words take *-de*, the majority of speakers choose *-de*. Ernestus & Baayen (2006) found that listeners also choose *-de* more often when the final obstruent is realized with weak voicing. This shows that listeners base their choice in addition on the detailed acoustic characteristics of the words. They are sensitive to incomplete neutralization even when this is not a requirement for the task that they are performing.

Lahiri, Jongman & Sereno (1990) also studied intraparadigmatic effects in the processing of voice. They carried out an experiment in which listeners of Dutch heard a verbal stem followed by the clitic pronoun *d'r* 'her' (the prime), and then performed auditory lexical decision on the same verbal stem in isolation (the target), realized with a voiceless final obstruent. The consonant cluster in the prime, consisting of the final obstruent of the verbal stem and the initial consonant of the clitic pronoun, was realized as voiceless (e.g., [kistər] *kies d'r* 'choose her'), or as voiced ([kizdər]). Both realizations are well-formed in Dutch (e.g., Zonneveld 1983). Listeners appeared to respond faster to target words ending in alternating obstruents when these obstruents were voiced in the prime, and to target words ending in non-alternating obstruents when these obstruents were voiceless in the prime. Unfortunately, Lahiri et al. do not report details on the acoustic characteristics of the target words, and it is therefore not clear to which extent the participants' behaviour might have been affected by incomplete neutralization. In addition, the study reports no statistics, and it is therefore not clear for which obstruents the observed differences are actually statistically significant, and even whether any of the reported differences reach significance.

In another study, Jongman, Sereno, Raaijmakers & Lahiri (1992) reported an experiment suggesting that the interpretation of a vowel as phonologically long or short depends on the alternating/non-alternating character of the following voiceless final obstruent. Listeners seem to attribute part of the length of a vowel to weak voicing if the following obstruent is alternating. The study does not report any statistics, but it does report for each participant the crossover boundaries (in ms) for each of the three studied word pairs, which differ in the alternating/non-alternating character of the final obstruents. Our analysis of these data show, unfortunately, that the differences between the word pairs fail to reach significance at the five-percent level (effect of word pair in an analysis of variance: $F(2, 42) = 2.70$; $p = 0.08$), so that it is uncertain what conclusions might be drawn from this study.

Summing up, previous studies have shown that alternating obstruents tend to be weakly voiced, at least in careful speech, and that listeners are sensitive to the subtle cues for weak voicing in the acoustic signal. In other words, the intraparadigmatic realizations of an obstruent affect its production in word-final position by inducing incomplete neutralization, and they affect word comprehension by the mediation of incomplete neutralization in the acoustic signal.

The present study addresses the question of whether voice alternation might affect speech perception over and above mediation by incomplete neutralization in

the acoustic signal. That is, does the listeners' knowledge about a word's morphological paradigm codetermine their percept of this word? Thus, does the lexical information about the plurals of [mant] and [krant] ([mɑndən] and [krɑntən]) cause listeners to perceive the [t] of [mant] as more voiced than the [t] of [krant]?

We report a transcription experiment in which listeners, who are not phonetically trained, rated word-final obstruents on a five point scale as voiceless or voiced. Several authors (e.g., Vieregge 1987, Cucchiariini 1993) have argued that the phonetic transcription of speech sounds may be affected by their orthographic representations, the phonotactics of the language, and semantics. In addition, Kemps, Ernestus, Schreuder & Baayen (2005) have shown that the transcription of reduced word forms is affected by their unreduced counterparts. Listeners report the presence of [l] in words ending in the suffix *-lijk* [-lək], even when this suffix is reduced to [k]. In this study, we investigated whether phonetic transcriptions made by naive participants might also be affected by lexical information about the segments' realizations in the word's morphological paradigm. Such information is of no use in a transcription task, where listeners have to base their judgments exclusively on the acoustic signal. In fact, intraparadigmatic effects, if present, would give rise to less accurate phonetic transcriptions. Hence, if we find such intraparadigmatic effects for naive participants, these effects must be automatic, that is, not available to participants for conscious, strategic control, and therefore a characteristic of everyday speech perception.

We opted for a rating task, instead of a traditional transcription task in which listeners represent sounds by IPA-symbols, because a rating may reveal more subtle differences in the listeners' percept of the voicing of alternating and non-alternating obstruents. We presented one group of listeners with full words, and another group with the final rimes of the same words. Since incomplete neutralization may be present in the acoustic signal, we may expect a difference between the ratings for the alternating and non-alternating obstruents by all listeners. Crucially, the listeners hearing the words could identify all words presented, and their judgments could therefore also be affected by lexical information about the intraparadigmatic realizations of the final obstruents. In contrast, the listeners hearing the rimes could identify the presented words only in a smaller number of cases, and a lexical effect on their judgments, if present, would necessarily be smaller. Hence, if intraparadigmatic realizations affect the perception of voice, we may expect a difference between the two groups of listeners, such that listeners hearing full words rate alternating obstruents as more voiced.

We expect intraparadigmatic effects to be smaller for fricatives than for plosives, since in Dutch the voiced-voiceless opposition is weaker for fricatives. Many speakers of Dutch tend to realize /z/ as [s], and even more speakers tend to realize /v/ as [f] in all positions in the word (e.g., Collins & Mees 1981:159; Gussenhoven & Bremmer 1983:57). Furthermore, the voicing of fricatives is highly predictable after vowels, as within words voiced fricatives are nearly always preceded by long vowels, and voiceless fricatives by short vowels. Finally, the difference between alternating and non-alternating obstruents is represented in orthography for plosives only. Final plosives that alternate in voice are always rep-

represented as voiced, and non-alternating plosives are always represented as voiceless. To give an example, the orthographic representations *mand* and *krant*, which are both realized with [t], show that the former morpheme is realized with [d] in the plural, while the latter morpheme is always realized with [t]. Fricatives, in contrast, are invariably represented as voiceless at the end of syllables. Their orthographic representations reveal nothing of their alternating/non-alternating character. Thus both [bas] ‘bass’ and [bas] ‘boss’ are orthographically transcribed with *s* (*bas*, *baas*), although the plural of [bas] is [bazən] *bazen* with a voiced [z]. In other words, whereas orthography reinforces the voiced-voiceless opposition for plosives, it does not do so for fricatives.

The participants rated the voicing of the final obstruents on a five point scale. In order to ensure that this five point scale would correspond with the full range of completely voiceless to completely voiced, we included realizations with completely voiced final obstruents in the experiment, which are unnatural in Dutch. Thus, the experiment also included realizations such as [pard] (from [part] ‘horse’) and [kaz] (from [kas] ‘cheese’).

The materials and the recording procedure are described in section 2. This section also reports acoustic analyses that we carried out in order to ascertain whether our speaker had realized alternating obstruents as weakly voiced, and in order to document the acoustic characteristics of the voiced final obstruents. In section 3, we present the actual rating experiment. Section 4 summarizes the findings and presents our conclusions. Moreover, it discusses how to incorporate our findings in the grammar of Dutch.

2. *Materials*

We selected 94 monosyllabic Dutch nouns, listed in the Appendix. Of these words, 30 end in alternating (e.g., *mand*), and 29 in non-alternating (e.g., *krant*) bilabial or alveolar plosives. The other 35 words end in labiodental or alveolar fricatives, of which 17 alternate in voice (e.g., *slaaf* ‘slave’ with the plural [slavən] *slaven*), and 18 are always voiceless (e.g., *bes* ‘berry’ with the plural [bəsən] *bessen*).

We created a list containing two orthographic representations for each word. The word was spelled with a voiced final obstruent in one representation and with a voiceless obstruent in the other representation. Which representation is correct according to the spelling conventions of Dutch depends on the manner of articulation of the obstruent (plosive, fricative) and its realization in the word’s morphological paradigm. The two versions of each word were presented right after each other, and the list thus started as follows: *baard*, *baart*, *baas*, *baaz*, *band*, *bant*, We asked a male speaker of Dutch, who makes a clear distinction between all voiced obstruents in Dutch and their voiceless counterparts, to record the words in the list. He was instructed to realize final obstruents as voiced, when they were represented as voiced, and to realize them as voiceless, when they were represented as voiceless. Although our speaker was not a phonetician, he did not need explicit instruction on how to realize word-final obstruents as voiced. Apparently, speakers of Dutch have ideas about how to realize word-final obstruents as

voiced, even though voiced final obstruents do not occur in their language. The words were recorded on a DAT (BASF master 94) in a soundproof room by means of a portable DAT-recorder Aiwa HD S100 and a Sony microphone ECM MS957. The recordings were stored as .wav files (sample rate: 48 KHz) on a computer by means of the speech analysis package *Praat* (Boersma 1996). Two phoneticians checked whether the final obstruents were realized as intended (voiced versus voiceless). If not, we asked our speaker to realize these words anew. In addition, we also asked our speaker to re-record words that he had realized with a schwa after the final obstruent.

We then carried out acoustic measurements in order to investigate whether our speaker had realized the final obstruents with incomplete neutralization, and how the voiced final obstruents differed from the voiceless ones. We first measured the durations of the vowels. We defined the beginning of the vowel as the beginning of a regular pattern in the wave form with the characteristics of the vowel, and the end of the vowel as the (sudden) end of this regular pattern. In addition, we measured the durations of the closures and release noises (the bursts plus the following periods of aspiration) of the final plosives, and the total durations of the final fricatives. We took the closure of the plosive to end at the sudden increase in amplitude at the beginning of the burst, and we assumed that the fricative and the release noise of the plosive end where the amplitude of the wave form is nearly identical to that of the background noise. Finally, we determined the proportion of the obstruent that was realized with vocal fold vibration. We assumed vocal fold vibration to be present if the waveform was periodic, the spectrogram contained a voice bar, and we could hear the vocal fold vibration in the acoustic signal.

The durations of vowels and final consonants in a word are, among others, affected by the presence and quality of extra consonants in the coda (Waals 1999). Since our data set contains only few words with complex codas, and since these words differ in the quality of the extra consonant, we restricted all our acoustic analyses to the words with simplex codas. We investigated by means of analyses of variance whether the durations were affected by the phonological length of the vowel (long versus short), the manner of the obstruent (plosive, fricative), the actual realization of this obstruent (voiced or voiceless), and its voice alternation (alternating, non-alternating).

For the duration of the vowel, we found significant main effects for phonological length ($F(111,101) = 654.09$; $p < 0.001$) and for the manner of the following final obstruent ($F(111,101) = 339.78$; $p < 0.001$). The phonologically long vowels were on average 99.2 ms longer than the phonologically short vowels, and the vowels preceding fricatives were on average 62.6 ms longer than the vowels preceding plosives. Furthermore, we observed an interaction between the phonological length of the vowel and the manner of the obstruent ($F(111,101) = 6.32$; $p = 0.014$). The difference between phonologically long and short vowels was less pronounced before plosives (81.5 ms) than before fricatives (101.9 ms). Finally, we observed an interaction between the manner of the obstruent and its actual realization ($F(12,101) = 4.30$; $p = 0.017$). Vowels preceding voiced fricatives were on average 16.0 ms longer than vowels preceding voiceless fricatives (see Figure

1, upper panel), whereas actual realization did not affect the duration of vowels preceding plosives (actual realization for plosives: $p > 0.1$). In many languages, voiced obstruents are preceded by longer vowels than voiceless obstruents. This is also the case for intervocalic obstruents in Dutch (Slis & Cohen 1969). In contrast, if a speaker of Dutch realizes word-final obstruents as voiced, apparently this does not necessarily affect the length of the preceding vowel.

The length of the final obstruent was affected by the manner of the obstruent ($F(11,109) = 108.35$; $p < 0.001$) and by its actual realization ($F(11,109) = 95.28$; $p < 0.001$). Fricatives were on average 52.2 ms longer than plosives, and voiced obstruents were on average 48.2 ms shorter than voiceless obstruents (see Figure 1, central panel). The length of a final obstruent may be a perceptual cue to its voicing.

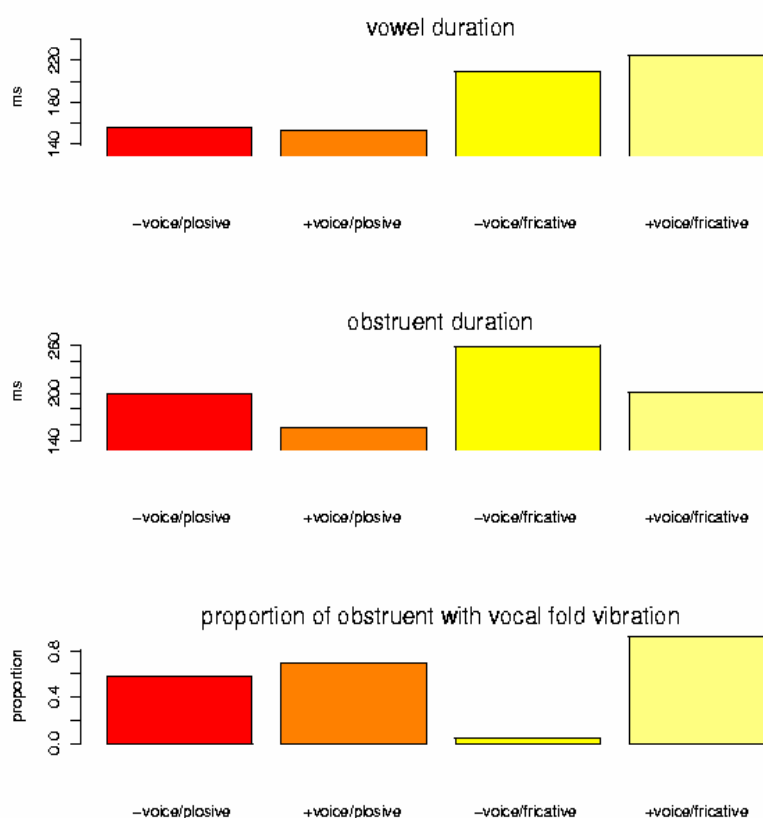


Figure 1: *The average duration of the vowel (upper panel), the average duration of the final obstruent (middle panel), and the average proportion of the final obstruent realized with vocal fold vibration (lower panel) for the words realized with voiced (+voice) and voiceless (-voice) final obstruents, broken for the manner of articulation of the final obstruent (plosive, fricative)*

We analysed separately the duration of the closure and the duration of the release noise for plosives. Closure duration was affected by the phonological length of the preceding vowel ($F(11,60) = 33.46$; $p < 0.001$), by the place of articulation

(bilabial or alveolar) of the plosive ($F(11,60) = 19.60$; $p < 0.001$), and its actual realization ($F(11,60) = 159.95$; $p < 0.001$). Closures following phonologically long vowels (75.1 ms) were on average shorter than closures following short vowels (94.9 ms), and alveolar closures were on average shorter (82.1 ms) than bilabial closures (104.4 ms). Interestingly, plosives that were actually voiced had longer closure durations (on average 108.9 ms) than plosives that were realized as voiceless (69.0 ms). This finding is in contrast with data for intervocalic positions, in which voiced plosives have shorter closures than voiceless plosives (e.g., Slis & Cohen 1969, Ernestus 2000). Probably, our speaker lengthened the voiced closures in order to have the presence of vocal fold vibration come out well. Vocal fold vibration is an important cue to voicing (see below), but preceding vowels mask the presence of vocal fold vibration in directly following closures, since they are relatively very loud. By lengthening the closures, our speaker made the presence of vocal fold vibration clearly audible. We also observed an interaction between the actual realization of the plosive and the phonological length of the preceding vowel ($F(11,60) = 8.94$; $p = 0.004$). Whereas the difference in closure duration between voiced and voiceless plosives was on average 33.6 ms after short vowels, it was 54.0 ms after long vowels. Finally, we observed an interaction between the actual realization of the plosive and its place of articulation ($F(11,60) = 15.45$; $p < 0.001$). The actual voice realization had a greater effect on the closures of alveolar plosives (average difference between voiced and voiceless alveolars: 49.8 ms) than on the closures of bilabial plosives (average difference: 16.7 ms).

The duration of the release noise was affected by the phonological length of the preceding vowel ($F(11,60) = 4.47$; $p = 0.039$) and by the actual realization of the plosive ($F(11,60) = 356.30$; $p < 0.001$). Release noises were longer after long vowels than after short vowels (on average 97.5 ms and 87.8 ms, respectively), and they were longer when the plosive was realized as voiceless than when it was voiced (on average 130.7 ms and 50.8 ms, respectively). The effect of the actual realization was larger (interaction between actual realization and place of articulation $F(1,60) = 16.58$; $p < 0.001$) for the alveolar plosives (the average release noise duration was 137.5 ms for [t]s, and 46.2 ms for [d]s) than for the bilabial plosives (average release noise duration for [p]s: 114.9 ms; for [b]s: 61.2 ms). Importantly, the duration of the release noise was also affected by the voice alternation of the plosive ($F(11,60) = 5.58$; $p < 0.001$). Release noises were shorter for alternating plosives (the average length was 84.7 ms) than for non-alternating plosives (93.7 ms). Our speaker realized the alternating plosives with weak voicing.

Finally, we analyzed the proportion of the obstruent that was realized with vocal fold vibration. The analysis of variance showed that the manner of the obstruent ($F(11,108) = 47.69$; $p < 0.001$) and its actual realization ($F(11,108) = 2568.41$; $p < 0.001$) affected the relative duration of vocal fold vibration in the obstruent. On average, 78.9% of the total duration of an actually voiced obstruent, and only 4.8% of a voiceless obstruent was realized with vocal fold vibration (see Figure 1, lower panel). This shows that our speaker realized final obstruents as voiced by keeping his vocal folds vibrating during a larger part of the obstruent. The differ-

ence between voiced and voiceless realizations was more pronounced for fricatives than plosives (interaction between actual realization and manner: $F(1,108) = 71.41$; $p < 0.001$). On average, both voiceless fricatives and voiceless plosives were realized with vocal fold vibration during less than 6.0% of their total duration (on average 3.5% and 5.8%, respectively), but voiced fricatives were realized with vocal fold vibration during 92.3%, and voiced plosives during 69.6% of their total duration.

In summary, our speaker signalled actual voicing especially by the presence of vocal fold vibration during a large part of the obstruent. In order to make the presence of vocal fold vibration well audible in plosives, he lengthened their closures. In addition, our speaker shortened voiced obstruents, and lengthened vowels preceding voiced fricatives (see also Figure 1). We found an effect of voice alternation only on the duration of the release noises, which suggests that our speaker only realized alternating plosives with weak voicing.

3. *The rating experiment*

3.1 *Method*

For the rating experiment, we divided the word tokens over two master lists. Each master list contained only one realization of each word type, that is, it contained either the realization with the voiced final obstruent (e.g., [krand]) or the realization with the voiceless final obstruent ([krant]). Furthermore, each master list contained both voiced and voiceless final obstruents, and both alternating and non-alternating obstruents. The words ending in the same type of obstruent (alveolar plosive, bilabial plosive, alveolar fricative, and labiodental fricative) were blocked in order to keep the options (e.g., [p] or [b] or something in between) in a sequence of trials constant. This facilitates the participants' task, which was difficult since the participants were not used to judging speech sounds. Moreover, presenting the words in blocks might enhance the probability that participants would differentiate between the different tokens of obstruents of the same type. We created ten versions of each master list by randomizing the words in the 17 blocks four times, while varying the order of the blocks. The resulting twenty lists are the lists with full words.

One group of participants listened to the full words, while another group listened to the rimes of the words starting at the steady states of the vowels. Thus, participants heard either [mɑnt], [hɑnd], etc., or [ɑnt], [ɔnd], etc. Since the rimes started at the steady state of the vowels, they contained no clear cues to the initial consonants, making it nearly impossible to trace the original words from which they had been spliced. Some of the initial consonants, however, might have been identifiable for some participants. In the worst case, this may have diminished the difference in scores between the participants hearing the full words and the participants hearing only the final rimes, which is the main interest of this study.

Vowels were shorter in the rimes, but since they all started at the steady states, the relative difference in duration between vowels preceding alternating and non-alternating obstruents was approximately the same in the words as in the rimes, as was also shown by a Linear Mixed Effect (LME) model (Pinheiro & Bates 2000,

Baayen, Tweedie & Schreuder 2002). This LME analysis had the duration of the vowel as its dependent variable, type of presentation (word, rime), phonological vowel length (long, short), the manner of the final obstruent (plosive, fricative), its voice alternation (alternating, non-alternating), and its actual realization (voiced, voiceless) as dependent variables, and word type as random effect variable. It revealed a main effect for the type of presentation ($F(11,165) = 16.34$; $p < 0.001$). Unsurprisingly, vowels were longer in the words than in the rimes. In addition, the model revealed main effects for the phonological length of the vowel ($F(11,165) = 626.53$; $p < 0.001$), the manner of the final obstruent ($F(11,52) = 305.99$; $p < 0.001$), and the actual realization of the final obstruent ($F(11,165) = 8.31$; $p = 0.005$). These variables affected vowel duration as described in section 2. Finally, the actual realization of the obstruent interacted with the type of the obstruent ($F(11,165) = 36.87$; $p < 0.001$), also as described in section 2. Importantly, the type of presentation did not interact with voice alternation. We may therefore assume that the acoustic cues for incomplete neutralization, including release noise duration and cues that we did not discover, are roughly similar for the obstruents in both words and rimes.

Some forty-six percent of the rimes (43 rimes) represented existing words of Dutch by themselves. For instance, the rime of [klet] *kleet* ‘carpet’ with the plural [kledən] represents the existing verb [et] *et* ‘eat’, which has the plural *eten* [etən]. The voice alternation for these words is independent of the voice alternation for the original words. We may therefore still expect that, if lexical intraparadigmatic information affects perception, the ratings for the full words will be more in conformity with the realizations of the obstruents in the words’ paradigms than the ratings for the rimes.

The participants were tested individually, sitting in a dimly lit room in front of a PC monitor and a panel with two buttons. They were asked to listen to the words that were presented to them, and to rate the final obstruents of the words that would be presented to them, depending on the block, as [b]s or [p]s, as [d]s or [t]s, as [v]s or [f]s, or as [z]s or [s]s, or as something in between. Participants received for each block a response form with a five-point scale for every trial. The grapheme for the voiceless variant of the obstruent illustrated the left most position of the scale, while the grapheme for the voiced variant illustrated the right-most position. In (1), we present a line from the response form for the rating of alveolar plosives.

(1) t 0 0 0 0 0 d

The course of a trial was as follows. The participant heard a warning beep of 377 Hz for 500 ms, followed by a pause of 200 ms. The participant then heard the stimulus, and rated the voicing of the final obstruent on the five point scale. The experiment was self-paced. Participants were presented with a new word or rime only after they had indicated that they were ready by pushing the right button.

Forty undergraduate students of the Radboud University Nijmegen were paid to participate in the experiment. Twenty students listened to the full words, while

the other twenty students listened to the rimes. The students were all native speakers of Dutch, and did not report any hearing deficits. Most of them originated from the southern part of the Netherlands, and may therefore be assumed to distinguish between the voiced and voiceless variants of both plosives and fricatives (Collins & Mees 1981:159, Gussenhoven & Bremmer 1983:57).

3.2 Results and discussion

Figure 2 presents the average scores by the participants who heard the final rimes (upper panel) and the participants who listened to the full words (lower panel). The left panels give the scores for the plosives, while the right panels list the scores for the fricatives. The scores are broken down for the realization of the final obstruent as intended by the speaker and the alternating/non-alternating character of the obstruent. A score of 1 indicates that the final obstruent was perceived as completely voiceless. A score of 5 indicates that the obstruent was completely voiced.

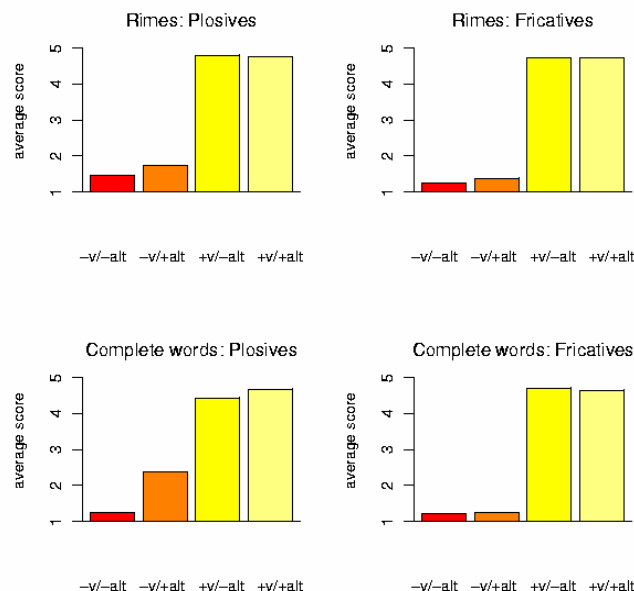


Figure 2: The average scores for the final obstruents by the participants who heard only the rimes (upper panel) and the participants who heard the words (lower panel). The left panels show the scores for the plosives, while the right panels show the scores for the fricatives. The scores are broken for the actual realization of the final obstruent as voiced or voiceless (+v or -v) and its voice alternation (+alt or -alt)

By means of a step-wise analysis of variance, we investigated whether the average scores for the items were affected by the type of presentation (rime, word), the actual realization of the final obstruent (voiced, voiceless), its voice alternation (alternating, non-alternating), and its manner of articulation (plosive, fricative). We removed four outlier stimuli (*hond* and *dood* both realized with [t], *schub* realized with [b], and *slurf* realized with [v]) from the data set in order to improve the normality of the model's residuals. The results are presented in Table 1.

Effects	DF	F-value	p-value
Manner	1	60.90	< 0.001
Alternation	1	87.58	< 0.001
Realization	1	12050.30	< 0.001
Manner: Alternation	1	40.02	< 0.001
Manner: Realization	1	64.12	< 0.001
Alternation: Realization	1	52.30	< 0.001
Alternation: Presentation	2	16.04	< 0.001
Realization: Presentation	1	23.03	< 0.001
Manner: Alternation: Realization	1	16.08	< 0.001
Manner: Alternation: Presentation	2	14.40	< 0.001
Manner: Realization: Presentation	1	14.42	< 0.001
Alternation: Realization: Presentation	1	8.49	0.004
Manner: Alternation: Realization: Presentation	1	6.63	0.010

Table 1: *Results of the stepwise analysis of variance of the average scores*

Figure 2 clearly shows that, unsurprisingly, by far the most important predictor for the average scores is the actual realization of the final obstruent. Whereas voiced obstruents were assigned an average score of 4.69, voiceless obstruents received an average score of only 1.54. This shows that listeners are sensitive to the acoustic cues of vowel duration, obstruent duration, and relative duration of vocal fold vibration for voicing.

Voice alternation also is an important factor (see Table 1). Scores were on average 0.28 higher for alternating obstruents than for non-alternating obstruents, but, as is clear from Figure 2, the effect is mainly carried by the voiceless plosives (as is supported by the interactions between voice alternation and manner of articulation, between voice alternation and actual realization, and between voice alternation, manner of articulation, and actual realization, as reported in Table 1). Whereas the average effect of voice alternation did not exceed 0.10 for the fricatives and the voiced plosives, it was 0.70 for the voiceless plosives. The focus of this study is on the interaction between voice alternation and type of presentation. Figure 2 suggests that an interaction is present for the voiceless plosives and the voiced plosives, but that it is absent for the fricatives. This is supported by separate analyses of variance on these four types of obstruents (see Table 2), as well as by the interactions between voice alternation and type of presentation with manner of articulation and actual realization in the overall analysis (Table 1). Voice alternation affects the rating for the voiceless plosives both in the rime condition ($F(11,55) = 6.99; p = 0.011$) and in the full word condition ($F(111,56) = 151.11; p < 0.001$), but the effect in the full word condition is larger (on average 1.11 in the full word condition versus 0.27 in the rime condition). The interaction between voice alternation and type of presentation is weaker for the voiced plosives than for the voiceless plosives (an analysis of just the plosives shows an interaction between voice alternation, type of presentation, and actual realization: $F(1,225) = 12.12; p < 0.001$), but again voice alternation has a larger effect on the full words

than on the rimes (an insignificant average difference of $-.04$ for the rimes and a significant difference of 0.24 for the full words). Since the words and the rimes do not differ in acoustic incomplete neutralization, the interaction of voice alternation with type of presentation must be due to lexical information that listeners access upon hearing words. The difference between words and rimes therefore shows that, at least for the final plosives in words, the effect of voice alternation is not only mediated by incomplete neutralization in the acoustic signal.

Type of obstruent	F-value	p-value
Voiceless plosives	38.04	< 0.001
Voiced plosives	10.00	0.002
Voiceless fricatives	0.70	> 0.1
Voiced fricatives	0.33	> 0.1

Table 2: *The interaction between voice alternation and type of presentation in separate analyses of variance of the average scores for the four types of final obstruents*

The interaction between voice alternation and type of presentation may be smaller for the voiced plosives than for the voiceless plosives, because the voiced realizations were unnatural for the listeners, who consequently scored nearly all voiced realizations as completely voiced. In addition, we may be observing a ceiling effect, because the voiced obstruents were already maximally voiced. The absence of an interaction for fricatives suggests that their voiced-voiceless opposition is too weak for our listeners, that is, there is not sufficient intraparadigmatic information for listeners to rely on for fricative final words.

The effects of the actual realization on the rating scores support the hypothesis that the ratings for the plosives in the full word condition were affected by lexical information. The obstruent's realization is expected to have a larger effect on the rating scores when listeners rely more on the acoustic signal. We find that while the difference between voiced and voiceless fricatives is approximately the same in the rime and in the full word condition, the difference in scores between voiced and voiceless plosives is smaller for words than for rimes (2.76 versus 3.20 ; see the interaction between actual realization, type of presentation, and manner of articulation given in Table 1). Apparently, listeners based their ratings less on the acoustic signal when they could identify the word (full word condition) and the obstruent was a plosive. In these cases, listeners based their ratings also on the paradigm of the word.

Our results hardly change if we restrict our analysis to the words of which the rimes are not existing words of Dutch by themselves. Such an analysis shows exactly the same main effects and interactions, except that the interaction between voice alternation, actual realization, and type of presentation is missing, probably due to the smaller number of data points. This shows that the effect of voice alternation on the scores for the rimes cannot be due only to the rimes that represent words by themselves, and thus to intraparadigmatic information that listeners ac-

cess upon hearing these rimes. The effect of voice alternation on the rimes is mainly due to incomplete neutralization in the signal.

In conclusion, alternating voiceless plosives were rated as more voiced than nonalternating voiceless plosives, especially when presented in full words. This finding shows that intraparadigmatic effects on listeners' voicing rates are mediated not only by incomplete neutralization in the acoustic signal, but also by listeners' paradigmatic knowledge.

4. *General discussion and conclusion*

This study addresses the question of whether intraparadigmatic effects in perception are mediated only by incomplete neutralization in the acoustic signal. We carried out a transcription experiment in which Dutch listeners who were not phonetically trained were presented with either full words or the final rimes of these same words, and were asked to rate the final obstruents as voiced or voiceless on a five point scale. Half of the words end in obstruents that are generally realized as voiceless, while the other half end in obstruents that alternate in voice, that is, are realized as voiced in some members of the words' paradigms.

The stimuli were recorded by a speaker of Dutch who realized final obstruents as voiced mainly by shortening them and realizing them with vocal fold vibration during a longer period. He lengthened preceding vowels only for fricatives. Interestingly, the cues that our speaker used to signal voicing were different from the cues that are used by speakers of English, for whom the duration of the preceding vowel is a major cue for all types of obstruents (e.g., Denes 1955). This difference may help explain the difficulties that speakers of Dutch experience with the voicing of word-final obstruents in English: Dutch speakers may focus on acoustic cues that are less relevant for English.

Acoustic analyses showed an effect of voice alternation only on the durations of the release noises of plosives. This is in line with the study by Ernestus and Baayen on pseudowords (2006), in which voice alternation was also found to affect only release noise duration. Possibly, we found no effect of voice alternation on vowel duration, in contrast to Warner et al. (2004), because the words with alternating and non-alternating obstruents in our experiment differed in their phonological make-up (i.e., they were not minimal pairs), and differences caused by the surrounding environment obscured any systematic vowel duration difference.

The main predictor for the average voicing scores by the listeners was the voice realization as intended by the speaker. In addition, scores were higher for alternating plosives than for non-alternating plosives. For the voiceless plosives, this was the case both in the rime and in the full word condition, while for the voiced plosives it was only so in the full word condition. Participants listening to rimes could not well identify the presented words in most cases, and their rating must therefore have been based mainly on the acoustic signal. We think that the intraparadigmatic effects on their scores are predominantly the result of incomplete neutralization in the acoustic signal, and therefore made possible by intraparadigmatic effects in production. In contrast, participants listening to full words could identify the presented words in all cases, and their scores could there-

fore be strongly affected by lexical knowledge as well. Indeed, the intraparadigmatic effects observed for the voiceless plosives were greater in the full word condition than in the rime condition. This shows that the knowledge of a word's paradigm also affects the interpretation of voicing. We conclude that intraparadigmatic effects in perception are partly mediated by the acoustic signal, and partly induced by the listener's lexical knowledge.

Note that our transcribers' intraparadigmatic knowledge in fact prevented them from basing their scores completely on the acoustic signal. Clearly, at least for non-trained transcribers, intraparadigmatic knowledge is in the way of phonetic transcriptions, intended as objective representations of the acoustic signal. Thus intraparadigmatic knowledge presents another problem for objective phonetic transcriptions, which Vieregge (1987) already claimed to be impossible. Furthermore, this finding also shows that intraparadigmatic effects mediated by listeners' lexical knowledge are automatic. They arise even when this is counterproductive for the task that listeners have to carry out. The intraparadigmatic effects in perception mediated by the listeners' knowledge may well be the motor behind incomplete neutralization in the acoustic signal in production. The alternating and non-alternating obstruents that were voiceless differed in their scores by only 0.21 in the rime condition but 0.73 in the full word condition. In other words, listeners' knowledge of the morphological paradigms give rise to an effect in the ratings that may be two times as big (0.73–0.21 versus 0.21) as that caused just by incomplete neutralization in the acoustic signal. The differences that listeners perceive between alternating and non-alternating obstruents is magnified by their knowledge of the paradigms, and this may encourage them to maintain these differences in their speech.

The intraparadigmatics of perception reported here instantiate in fact a subtle form of paradigmatic levelling. Steriade (2000) claims that paradigmatic levelling may affect phonemes as well as subphonemic characteristics. What we have shown here is that subphonemic paradigmatic levelling is not restricted to production, but is also a characteristic of perception.

Only the scores for final plosives were clearly affected by the words' paradigms. Recall that we found no evidence for incomplete neutralization in the signal for fricative-final words. This may explain why the alternating character of final fricatives did not affect the scores by the participants who heard just the final rimes. The fact that also the participants listening to the full words showed no effects, even though they could rely on their intraparadigmatic knowledge, may be explained by the weakness of the voiced-voiceless opposition for fricatives in Dutch. This opposition is not well maintained by most speakers of Dutch, it is hardly distinctive after vowels, and it is not supported by the spelling conventions of Dutch in syllable-final position.

Our listeners scored those voiceless obstruents as slightly voiced that are spelled as voiced. Nevertheless, the intraparadigmatic effects reported here cannot only be due to orthography, since our participants also showed sensitivity to incomplete neutralization in the signal when listening to rimes that do not represent words by themselves, and for which they consequently could not rely on the spell-

ing. Moreover, also other studies show that intraparadigmatic effects on the production and comprehension of voicing can be independent of orthography. As already mentioned in section 1, Dinnsen & Charles-Luce (1984) and Charles-Luce (1993) have shown that incomplete neutralization is also present in Catalan minimal word pairs for which the spelling does not reflect the difference between alternating and non-alternating obstruents. Furthermore, we have shown in a previous study that Dutch listeners are sensitive to weak voicing in fricatives, even though voiceless fricatives are always spelled as voiceless (Ernestus & Baayen 2006). Finally, we also have evidence that intraparadigmatic effects that are not mediated by incomplete neutralization do not just result from orthography. In a follow-up study (Ernestus & Baayen, in press), we presented the stimuli from the present rating experiment in a lexical decision experiment. We found that the frequency with which an alternating plosive is realized as voiced relative to the frequency with which it is voiceless affects response latencies. This frequency effect cannot be due to orthography, since both the voiced and voiceless realizations for alternating plosives are spelled as voiced. Intraparadigmatic effects need not be supported by orthography.

We now turn to the question of how to incorporate our findings in the grammar of Dutch. In generative grammar, intraparadigmatic effects are traditionally accounted for by means of underlying forms. A morpheme-final obstruent that is voiced before vowel-initial suffixes is also voiced in the underlying form, as already mentioned in the introduction of this paper. Thus, the underlying form of [mant] is /mand/ because of the voiced [d] in the plural [mɑndən]. Incomplete neutralization can be accounted for by the assumption that the voicing of the final obstruent in the underlying form affects the production of the obstruent via phonetic implementation rules preceding or coinciding with Final Devoicing, or replacing Final Devoicing (Dinnsen & Charles-Luce 1984, Port & O'Dell 1985, Slowiaczek & Dinnsen 1985).

A generative account of the data with a rule (or constraint) of Final Devoicing supplemented with various independent phonetic implementation rules (or constraints) is possible, but cumbersome. What is required to account for the data is (1) information as to whether a word-final (voiceless) obstruent is alternating, (2) Final Devoicing, and (3) phonetic implementation rules (constraints) that weakly (re)voice voiceless obstruents. By Occam's razor, we prefer a theory in which phonetic realization rules directly produce the correct form from the alternation information in the lexicon to a theory that devoices underlyingly voiced segments while partly re-voicing them in another stage in the derivation. In the OT-framework, we prefer a theory in which the constraint of Final Devoicing is simply dispensed with.

This line of argument can be taken a step further, since the mental lexicon contains form representations for a great many words, including inflected forms (e.g., Jackendoff 1975, Baayen, Dijkstra & Schreuder 1997, Alegre & Gordon 1999, Baayen, McQueen, Dijkstra & Schreuder 2003). Given that nearly every word is lexically represented, we may assume lexical representations that directly reflect the words' pronunciations. That is, the singular *mand* may be represented as

/mant/ and the plural *manden* as /mɑndən/. As a consequence, all theoretical approaches based on lexical representations that do not reflect actual pronunciation are unnecessarily complex, irrespective of whether they assume Final Devoicing. Hence, by Occam's razor, we also disprefer accounts in which phonetic implementation rules are directly applied to underlying voiced obstruents or to archiphonemes unspecified for voice (cf. Trubetzkoy 1958, Lieb 1998), even though these accounts can easily incorporate incomplete neutralization.

Following Bybee (2001, see also Ernestus & Baayen 2006), we view intraparadigmatic effects in production and perception as resulting from lexical analogy that is not mediated by abstract underlying representations nor by an idiosyncratic feature that marks a final obstruent as alternating or non-alternating. When speakers realize [mant] in production, they activate [mant] as well as [mɑndən]. The plural contains [d], which may affect the realization of the word-final [t], resulting in weak voicing. In contrast, when speakers realize [krant], the inflectional paradigm members do not contain [d], and they do not slightly voice the word-final obstruent.

Lexical analogy can also explain intraparadigmatic effects in perception. When listeners perceive a word, this word as well as the morphologically related words are activated in the listeners' mental lexicon, and these paradigmatic competitors codetermine the listeners' percept. Formalized analogical models, such as Skousen's Analogical Model of Language (Skousen 1989, 1993), can easily incorporate these intraparadigmatic effects.

One of the predictions of our approach, that we leave for further research, is that alternating intervocalic voiced obstruents would be realized and perceived as less voiced than their non-alternating counterparts. For instance, the alternating [d] of Dutch [bɪdən] *bidden* 'to pray' (with the singular present-tense [bit] *bid*) would be produced and perceived as less voiced than the non-alternating [d] of [mɪdən] *midden* 'middle'. Note that under a generative account with abstract underlying representations, we would not expect a difference between [bɪdən] and [mɪdən], since both forms underlyingly contain /d/ (/bɪdən/, /mɪdən/). Thus, alternating and non-alternating voiced intervocalic obstruents would form a good test case for the generative account and the lexical paradigmatic account.

To conclude, the main point of this study is that the intraparadigmatic realization of an obstruent does not only affect its production but also its perception. The intraparadigmatic effects in perception are partly mediated by incomplete neutralization in the acoustic signal, and partly arise due to listeners' lexical knowledge. These intraparadigmatic effects are automatic, and prevent listeners from producing accurate phonetic transcriptions that are true reflections of the acoustic signal itself.

Acknowledgements

We would like to thank Louis Pols, Erik Jan van der Torre, Jeroen van de Weijer, and an anonymous reviewer for their helpful comments on an earlier version of this paper.

References

- Alegre, Maria & Peter Gordon. 1999. "Frequency Effects and the Representational Status of Regular Inflections". *Journal of Memory and Language* 40.41-61.
- Baayen, R. Harald, Ton Dijkstra & Rob Schreuder. 1997. "Singulars and Plurals in Dutch: Evidence for a Parallel Dual Route Model". *Journal of Memory and Language* 37.94-117.
- Baayen, R. Harald, James McQueen, Ton Dijkstra & Rob Schreuder. 2003. "Frequency Effects in Regular Inflectional Morphology: Revisiting Dutch Plurals". *Morphological Structure in Language Processing* ed. by R. Harald Baayen & Rob Schreuder, 355-390. Berlin: Mouton de Gruyter.
- Baayen, R. Harald, Fiona J. Tweedie & Rob Schreuder. 2002. "The Subjects as a Simple Random Effect Fallacy: Subject variability and Morphological Family Effects in the Mental Lexicon". *Brain and Language* 81.55-65.
- Boersma, Paul. 1996. Praat: Doing Phonetics by Computer. Ms., University of Amsterdam.
- Booij, Geert E. 1995. *The Phonology of Dutch*. Oxford: Clarendon Press.
- Bybee, Joan L. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Charles-Luce, Jan. 1993. "The Effects of Semantic Context on Voicing Neutralization". *Phonetica* 50.28-43.
- Collins, Beverley S. & Inger Mees. 1981. *The Sounds of English and Dutch*. Leiden: Leiden University Press.
- Cucchiariini, Catia. 1993. Phonetic Transcription: A Methodological and Empirical Study. Ms., University of Nijmegen.
- Denes, Paul. 1955. "Effect of Duration on the Perception of Voicing". *Journal of the Acoustical Society of America* 27.761-764.
- Dinnsen, Daniel A. & Jan Charles-Luce. 1984. "Phonological Neutralization, Phonetic Implementation and Individual Differences". *Journal of Phonetics* 12.49-60.
- Ernestus, Mirjam. 2000. *Voice Assimilation and Segment Reduction in Casual Dutch. A Corpus-Based Study of the Phonology-Phonetics Interface*. Ph.D. dissertation, Vrije Universiteit Amsterdam.
- Ernestus, Mirjam & R. Harald Baayen. 2003. "Predicting the Unpredictable: Interpreting Neutralized Segments in Dutch". *Language* 79.5-38.
- Ernestus, Mirjam & R. Harald Baayen. 2006. "The Functionality of Incomplete Neutralization in Dutch: The Case of Past-Tense Formation". *Laboratory Phonology 8* ed. by Louis Goldstein, D.H. Whalen & Catherine T. Best, 27-49. Berlin: Mouton de Gruyter.
- Ernestus, Mirjam & R. Harald Baayen. In press. "Paradigmatic Effects in Auditory Word Recognition: The Case of Alternating Voice in Dutch". *Language and Cognitive Processes*.
- Fourakis, Marios & Gregory K. Iverson. 1984. "On the 'Incomplete Neutralization' of German Final Obstruents". *Phonetica* 41.140-149.

- Gussenhoven, Carlos & Rolf H. Bremmer Jr. 1983. "Voiced Fricatives in Dutch: Sources and Present-Day Usage". *North-Western European Language Evolution* 2.55-71.
- Jackendoff, Ray S. 1975. "Morphological and Semantic Regularities in the Lexicon". *Language* 51.639-671.
- Jongman, Allard, Joan A. Sereno, Marianne Raaijmakers & Aditi Lahiri. 1992. "The Phonological Representation of [Voice] in Speech Perception". *Language and Speech* 35.137-152.
- Kemps, Rachel, Mirjam Ernestus, Rob Schreuder & R. Harald Baayen. 2005. "Prosodic Cues for Morphological Complexity: The Case of Dutch Noun Plurals". *Memory and Cognition* 33.430-446.
- Lahiri, Aditi, Allard Jongman & Joan A. Sereno. 1990. "The Pronominal Clitic *d'r* in Dutch: A Theoretical and Experimental approach". *Yearbook of Morphology* 3.115-127.
- Lieb, Hans-Heinrich. 1998. "Morph, Wort, Silbe: Umrisse einer Integrativen Phonologie des Deutschen". *Variation und Stabilität in der Wortstruktur: Untersuchungen zu Entwicklung, Erwerb und Varietäten des Deutschen und anderer Sprachen* ed. by Matthias Butt & Nanna Fuhrhop (= *Germanistische Linguistik* 141-142), 334-407. Hildesheim: Georg Olms Verlag.
- Pinheiro, José C. & Douglas M. Bates. 2000. *Mixed-Effects Models in S and S-PLUS*. New York: Statistics and Computing, Springer.
- Port, Robert & Penny Crawford. 1989. "Incomplete Neutralization and Pragmatics in German". *Journal of Phonetics* 17.257-282.
- Port, Robert & Michael O'Dell. 1985. "Neutralization of Syllable-Final Voicing in German". *Journal of Phonetics* 13.455-471.
- Skousen, Royal. 1989. *Analogical Modeling of Language*. Dordrecht: Kluwer.
- Skousen, Royal. 1993. *Analogy and Structure*. Dordrecht: Kluwer.
- Slis, Iman H. & Antonie Cohen. 1969. "On the Complex Regulating the Voiced-Voiceless Distinction". *Language and Speech* 12.80-102; 137-155.
- Slowiaczek, Louisa & Daniel A. Dinnsen. 1985. "On the Neutralizing Status of Polish Word-Final Devoicing". *Journal of Phonetics* 13.325-341.
- Steriade, Donca. 2000. "Paradigm Uniformity and the Phonetics-Phonology Boundary". *Papers in Laboratory Phonology V: Language Acquisition and the lexicon* ed. by Michael Broe & Janet Pierrehumbert, 313-334. Cambridge: Cambridge University Press.
- Trubetzkoy, Nikolai S. 1958. *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.
- Vieregge, Wilhelm. 1987. "Basic Aspects of Phonetic Segmental Transcription". *Probleme der Phonetischen Transkription* ed. by Antonio Almeida & Angelika Braun, 5-55. Stuttgart: Franz Steiner Verlag Wiesbaden.
- Waals, Juliette. 1999. *An Experimental View of the Dutch Syllable*. The Hague: Holland Academic Graphics.
- Warner, Natasha, Erin Good, Allard Jongman & Joan A. Sereno. 2006. "Orthographic vs. Morphological Incomplete Neutralization Effects", *Journal of Phonetics* 34.285-293.

- Warner, Natasha, Allard Jongman, Joan A. Sereno & Rachel Kemps. 2004. "Incomplete Neutralization and Other Sub-Phonemic Durational Differences in Production and Perception: Evidence from Dutch". *Journal of Phonetics* 32.251-276.
- Zonneveld, Wim. 1983. "Lexical and Phonological Properties of Dutch Devoicing Assimilation". *Sound Structures: Studies for Antonie Cohen* ed. by Marcel van den Broecke, Vincent J. van Heuven & Wim Zonneveld, 297-312. Dordrecht: Foris.

Appendix

- Experimental words ending in obstruents that are voiced in inflectionally related words:

<i>krib</i>	'manger'	<i>kleed</i>	'cloth'	<i>grens</i>	'border'
<i>kwab</i>	'lobe'	<i>koord</i>	'cord'	<i>hals</i>	'neck'
<i>rib</i>	'rib'	<i>maand</i>	'month'	<i>kaas</i>	'cheese'
<i>schub</i>	'scale'	<i>mand</i>	'basket'	<i>laars</i>	'boot'
<i>web</i>	'web'	<i>moord</i>	'murder'	<i>muis</i>	'mouse'
<i>baard</i>	'beard'	<i>naald</i>	'needle'	<i>neus</i>	'nose'
<i>bed</i>	'bed'	<i>oord</i>	'place'	<i>prijs</i>	'price'
<i>brand</i>	'fire'	<i>paard</i>	'horse'	<i>spijs</i>	'food'
<i>brood</i>	'bread'	<i>strand</i>	'beach'	<i>korf</i>	'basket'
<i>bruid</i>	'bride'	<i>tand</i>	'tooth'	<i>scherf</i>	'fragment'
<i>dood</i>	'dead'	<i>veld</i>	'field'	<i>slaaf</i>	'slave'
<i>eend</i>	'duck'	<i>vod</i>	'rag'	<i>slurf</i>	'trunk'
<i>hand</i>	'hand'	<i>woord</i>	'word'	<i>staaf</i>	'bar'
<i>held</i>	'hero'	<i>zwaard</i>	'sword'	<i>wolf</i>	'wolf'
<i>hemd</i>	'shirt'	<i>baas</i>	'boss'	<i>zalf</i>	'ointment'
<i>hond</i>	'dog'	<i>gans</i>	'goose'		

- Experimental words ending in obstruents that are always voiceless:

<i>klap</i>	'bang'	<i>lat</i>	'slat'	<i>fles</i>	'bottle'
<i>mep</i>	'clout'	<i>lint</i>	'ribbon'	<i>kous</i>	'stocking'
<i>schep</i>	'scoop'	<i>maat</i>	'measure'	<i>pols</i>	'wrist'
<i>stip</i>	'dot'	<i>mot</i>	'moth'	<i>tas</i>	'bag'
<i>strip</i>	'strip'	<i>pet</i>	'cap'	<i>tros</i>	'cluster'
<i>beurt</i>	'turn'	<i>pit</i>	'pip'	<i>vis</i>	'fish'
<i>cent</i>	'cent'	<i>poort</i>	'gate'	<i>zeis</i>	'scythe'
<i>fluit</i>	'flute'	<i>put</i>	'well'	<i>bef</i>	'jabot'
<i>geit</i>	'goat'	<i>schat</i>	'treasure'	<i>juf</i>	'female teacher'
<i>grot</i>	'cave'	<i>scheut</i>	'twinge'		
<i>hert</i>	'deer'	<i>spruit</i>	'sprout'	<i>nimf</i>	'nymph'
<i>kat</i>	'cat'	<i>staart</i>	'tail'	<i>plof</i>	'thud'
<i>klant</i>	'customer'	<i>bes</i>	'berry'	<i>rif</i>	'reef'
<i>knot</i>	'knot'	<i>bos</i>	'woods'	<i>slof</i>	'slipper'
<i>krat</i>	'crate'	<i>dans</i>	'dance'	<i>straf</i>	'punish- ment'
<i>kreet</i>	'cry'	<i>eis</i>	'require- ment'		
<i>krot</i>	'slum'				