# Paradigmatic effects in auditory word recognition : The case of alternating voice in Dutch — Source link ↗

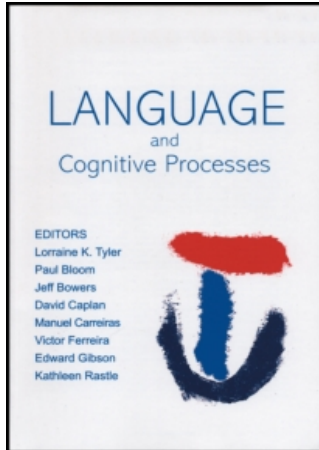Mirjam Ernestus, R. Harald Baayen

**Institutions:** Max Planck Society

Related papers:

- The locus of the orthographic consistency effect in auditory word recognition

- Orthography shapes the perception of speech: The consistency effect in auditory word recognition

- Orthographic effects on rhyme monitoring.

- The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French

- Praat : doing phonetics by computer

# Language and Cognitive Processes

## Paradigmatic effects in auditory word recognition: The case of alternating voice in Dutch

Mirjam Ernestus [a]; Harald Baayen [a]
[a] Max-Planck Institute for Psycholinguistics and Radboud University Nijmegen, Nijmegen, The Netherlands

PLEASE SCROLL DOWN FOR ARTICLE

Ψ **Psychology Press**
Taylor & Francis Group

# Paradigmatic effects in auditory word recognition: The case of alternating voice in Dutch

Mirjam Ernestus and Harald Baayen

*Max-Planck Institute for Psycholinguistics and Radboud University Nijmegen, Nijmegen, The Netherlands*

Two lexical decision experiments addressed the role of paradigmatic effects in auditory word recognition. Experiment 1 showed that listeners classified a form with an incorrectly voiced final obstruent more readily as a word if the obstruent is realised as voiced in other forms of that word's morphological paradigm. Moreover, if such was the case, the exact probability of paradigmatic voicing emerged as a significant predictor of the response latencies. A greater probability of voicing correlated with longer response latencies for words correctly realised with voiceless final obstruents. A similar effect of this probability was observed in Experiment 2 for words with completely voiceless or weakly voiced (incompletely neutralised) final obstruents. These data demonstrate the relevance of paradigmatically related complex words for the processing of morphologically simple words in auditory word recognition.

## INTRODUCTION

Auditory word recognition involves the activation of multiple lexical representations (e.g., Goldinger, Luce, & Pisoni, 1989; Pisoni, Nusbaum, Luce, & Slowiaczek, 1985). Phonological neighbours compete in the recognition process, and the speed with which a spoken word is recognised is affected by the density of its phonological neighbourhood, as well as by the frequencies of occurrence of these neighbours. In white noise, listeners recognise those words more easily that have fewer phonological neighbours with a higher frequency (Luce, 1985).

Recent research has shown that the mental lexicon not only contains monomorphemic words and morphologically complex words with unpredictable phonological, morphological, syntactic, or semantic characteristics. Fully regular morphologically complex words, including inflectional word forms, also leave traces in lexical memory. Thus, several studies have shown surface frequency effects in the comprehension of fully regular inflections in the visual modality for Dutch (Baayen, Dijkstra, & Schreuder, 1997; Baayen, Schreuder, De Jong, & Krott, 2002; Schreuder, De Jong, Krott, & Baayen, 1999), English (Alegre & Gordon, 1999; Sereno & Jongman, 1997; Taft, 1979), Finnish (Bertram, Laine, Baayen, Schreuder, & Hyönä, 1999), and Italian (Baayen, Burani, & Schreuder, 1997), and in the auditory modality for Dutch (Baayen, McQueen, Dijkstra, & Schreuder, 2003).

Since inflectionally related forms often share their initial phonemes (cf. English singular <u>book</u> and plural <u>books</u>), the storage of fully inflected forms may lead to paradigmatic competition: The recognition of one word form may be hampered by the presence of inflectionally related forms in the mental lexicon—effectively phonological neighbours—and especially so by those with relatively high frequencies. Whereas several cross-model priming studies have shown that the recognition of a morphologically complex word is affected by words sharing the same stem (e.g., Boudelaa & Marslen-Wilson, 2004; Emmorey, 1989; Marslen-Wilson, Tyler, Waksler & Older, 1994; Meunier & Segui, 2001; Reid & Marslen-Wilson, 2003), the possibility of competition between inflectional variants in unprimed word recognition has received little attention in the literature.

One of the few exceptions is a study by Kemps, Ernestus, Schreuder, and Baayen (2005), which suggests that there is such competition (see also Baayen et al., 2003), and that it is attenuated by subphonemic cues. The authors investigated the processing of singular and plural noun forms in Dutch. Dutch plural nouns, as their English counterparts, consist of the noun stem plus a suffix, so that the singular is onset embedded in the plural. Previous experimental work by Davis, Marslen-Wilson, and Gaskell (2002) and by Salverda, Dalian, and McQueen (2003) has shown that onset embedded words tend to be shorter in their carrier words than when they are realised in isolation (e.g., the syllable *ham* in *hamster* is shorter than the word *ham*), and that listeners take advantage of this durational difference. Kemps et al. showed that this is also true for words that are onset-embedded in words belonging to the same inflectional paradigm. In the presence of the plural suffix, the stem is shorter: the singular and plural forms differ with respect to the durations of the segments that they have in common. Thus the sequence [buk] "book" is longer in the singular <u>boek</u> [buk] than in the plural <u>boeken</u> [bukə]. Using the cross-splicing technique, they also showed that a mismatch between segmental and durational cues leads to delayed responses in number decision and auditory lexical decision. Listeners reacted faster to

[buk] with the normal durational pattern of a singular, and to [bukə] with the normal durational pattern of a plural than to these same forms with the durational patterns of the opposite number.

The present study further investigates paradigmatic competition in auditory word recognition. The test case are obstruent-final words in Dutch. Morpheme-final obstruents in Dutch may be voiced or voiceless, as illustrated by minimal pairs such as [notə] (noot+en) "nuts" and [nodə] (nood+en) "necessities", which are both plural nouns that consist of a noun stem and the plural suffix [ə]. In syllable-final position, however, Dutch obstruents are voiceless (except before voiced plosives which may induce regressive voice assimilation). Hence, the singular of both [notə] and [nodə] is [not].

The paradigmatic alternation of voiced and voiceless obstruents (e.g., [nodə] "necessities" vs. [not] "necessity") is widespread in the lexicon of Dutch and affects the fine acoustic details of the voiceless realisation of the obstruent in word-final position. Although all word-final obstruents are voiceless, the alternating obstruents tend to have more acoustic character-istics of voiced obstruents than non-alternating obstruents, which are always voiceless (see for Dutch, Warner, Jongman, Sereno, & Kemps, 2004; Ernestus & Baayen, in press a, in press b; and for the effect in other languages with final devoicing, e.g., Charles-Luce, 1993; Dinnsen & Charles-Luce, 1984; Port & Crawford, 1989; Port & O'Dell, 1985; Slowiaczek & Dinnsen, 1985). Word-final alternating obstruents tend to be shorter. They tend to be realised with vocal fold vibration during a longer period, and they are generally preceded by longer vowels. Thus, the [t] of [not] "necessity" (plural [nodə]) tends to have more acoustic characteristics of voiced obstruents than the [t] of [not] "nut" (plural [notə]). In other words, the neutralisation of voice at word-final position is incomplete. In what follows, we will refer to voiceless obstruents that possess some acoustic characteristics of genuine voiced obstruents as weakly voiced.

Listeners take advantage of the acoustic differences between alternating and non-alternating obstruents, even though these differences are very subtle. Listeners are able to infer at above chance level the correct spelling for 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 & Crawford, 1989; Port & O'Dell, 1985; Warner et al, 2004). Thus, when Dutch listeners hear [not], they are more likely to report nood when the final obstruent is weakly voiced, and noot when the obstruent is completely voiceless.

Listeners also show sensitivity to incomplete neutralisation in a task that does not force them to use these subtle subsegmental cues. 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 obstruent. If the

obstruent is always realised as voiceless, the appropriate allomorph is -te, otherwise it is -de. Ernestus and Baayen (2003) presented listeners with the stems of pseudo-verbs ending in obstruents, which, being word-final, were necessarily realised as voiceless. The listeners were asked to write down the corresponding past tense forms, but, as the presented verbs were nonexistent, they had no information about whether the final obstruent was alternating. It was observed that participants tended to base their choice between -te and -de on the word's phonological similarity neighbourhood consisting of the words ending in the same type of rime. If most words in the phonological neighbourhood ended in alternating obstruents, speakers tended to choose -de, and if most words ended in non-alternating obstruents, the majority of speakers chose -te. Ernestus and Baayen (in press a) found that listeners also choose -de more often if the final obstruent is realised with weak voicing. This shows that the choice between -de and -te is also affected by the detailed acoustic characteristics of the words. Listeners show sensitivity to incomplete neutralisation, even when this is not strictly necessary for the task that they are performing.

The paradigmatic effects of voice alternation in speech perception are not restricted to effects mediated by incomplete neutralisation in the acoustic signal. Ernestus and Baayen (in press b) asked two groups of listeners to rate word-final obstruents as voiced or voiceless on a five-point scale. One group of listeners heard complete words. They scored alternating voiceless obstruents as more voiced than the second group of participants, who heard just the final rimes of the same words. For instance, participants hearing the full word [krɪp] "manger" (plural [krɪbə]) scored the final [p] as more voiced than the participants who heard just [ɪp] (a non-word). Apparently, the percept of an obstruent is affected by how this obstruent is realised in the inflectional variants of its word's paradigm.

In the present paper, we investigated the effect of paradigmatic voice alternation on spoken word recognition. In Experiment 1, a lexical decision experiment, we presented listeners with singulars that were realised either correctly with a voiceless word-final obstruent, such as [hɑnt] "hand" and [krɑnt] "newspaper", or incorrectly with a voiced final obstruent, "[hɑnd] and *[krɑnd]. Here and in what follows, we mark realisations of words with incorrectly voiced final obstruents with an asterisk. Half of the noun stems ended in alternating obstruents, such as [hɑnt] with the plural [hɑndə], and an incorrect voiced realisation in the singular (*[hɑnd]) was therefore supported by the word's inflectional paradigm, where we find [hɑndə]. The other half of the words ended in non-alternating obstruents. Their paradigms did not support an incorrect voiced realisation of the singular-final obstruent. For instance, the plural of [krɑnt] is [krɑntə], which does not support the final [d] of *[krɑnd]. We asked participants to say "yes" to forms such as *[hɑnd]. We pointed out to the participants that these realisations

can occur in compounds (as the result of regressive voice assimilation). For instance, *[hɑnd] occurs in the compound [hɑndbuk] "hand book". We expected listeners to be hampered less by an incorrect voiced final obstruent if this obstruent has voiced allophones in the word's inflectional paradigm. That is, the participants should show fewer errors and shorter response latencies for those incorrect forms that have paradigmatic support for voicing.

In addition, we expected that voiced allophones have larger effects on the comprehension of plosives than on fricatives. Voiced plosives are nearly always realised as voiced, by all speakers of Dutch. Only the initial /d/-s of some function words may be realised as voiceless after obstruents. Voiced fricatives, on the contrary, are often realised as voiceless. They are systematically devoiced after obstruents (e.g., Booij, 1995), they tend to be devoiced in utterance initial position, and speakers from large parts of the Netherlands even realise all fricatives as voiceless (e.g., Collins & Mees, 1981, p. 159; Gussenhoven & Bremmer, 1983, p. 57). Moreover, Ernestus and Baayen (in press b) showed that the rating of final voiceless plosives as voiced or voiceless on a five-point scale is affected by their alternating character, whereas this is not the case for voiceless fricatives.

In Experiment 1, we also addressed the role of the frequencies of occurrence of the paradigmatic competitors, and in Experiment 2, their possible interaction with subphonemic cues. We postpone the discussion of these frequency effects until we have discussed the effects of the categorical measure indicating whether a voiced realisation does or does not receive paradigmatic support.

## EXPERIMENT 1

### Method

*Participants.* Forty-one native speakers of Dutch were paid to participate in the experiment. Most of them were students at the Radboud University Nijmegen.

*Materials.* The materials comprised 94 existing Dutch word types, listed in the Appendix. Of these words, 30 ended in alternating, and 29 in non-alternating bilabial or alveolar plosives, for instance, hand ([hɑnt]) "hand" (plural handen [hɑndə]) with an alternating plosive, and krant ([krɑnt]) "newspaper" (plural kranten [krɑntə]) with a non-alternating plosive. The other 35 words ended in labiodental or alveolar fricatives of which 17 alternated in voice and 18 were always voiceless, for instance, alternating baas ([bas]) "boss" (plural bazen [bazə]) and non-alternating bos ([bɔs]) "woods" (plural bossen [bɔsə]). We included two tokens for every word:

In one token the final obstruent was correctly realised as voiceless (e.g., [hɑnt]), in the other token it was incorrectly voiced (*[hɑnd]). A rating study with 10 participants (Ernestus & Baayen, in press b) showed that the final obstruents intended as voiceless and those intended as voiced indeed sounded as such, The correct voiceless realisations of alternating voiceless plosives showed weak voicing: Their release noises were on average 9 ms shorter than the release noises of the non-alternating voiceless plosives, which led to slightly higher voicing scores by the participants in the rating experiment.

In order to have the words in three different orders, we created three master lists with their complementary lists. The master lists contained one token of every word type, 24 or 25 tokens ending in a voiceless and 24 or 25 tokens ending in a voiced alveolar plosive ([t] and [d]), 5 tokens ending in a voiceless and 5 tokens ending in a voiced bilabial plosive ([p] and [b]), 10 or 11 words ending in a voiceless and 10 or 11 words ending in a voiced alveolar fricative ([s] and [z]), and, finally, 7 words ending in a voiceless and 7 words ending in a voiced labiodental fricative ([v] and [f]). The complementary lists contained the tokens that were not incorporated in the master lists, such that a master list and its complementary list together contained both tokens of every type. We added 358 filler words to each list. We randomised the order of the words in the master lists and adapted the orders in the corresponding complementary lists such that a word occupied the same position in a master list and in the corresponding complementary list. Each experimental list was preceded by 16 practice items.

*Procedure.* Participants were instructed to decide as quickly as possible whether the form they heard was a word or a pseudoword. They were explicitly instructed to classify also as words those items that were existing words incorrectly realised with voiced final obstruents, such as *[hɑnd] and *[krɑnd]. As mentioned above, these realisations may occur in compounds, and are only illegal in isolation. Participants responded by pressing the "yes" or "no" button on a button box. Each trial consisted of the presentation of a warning tone (377 Hz) of 500 ms, followed after an interval of 450 ms by the auditory stimulus. Stimuli were presented through Sennheiser headphones. Reaction times were measured from stimulus offset. Each new trial was initiated 2500 ms after offset of the previous stimulus. When a participant did not respond within 2500 ms post-offset, a time-out response was recorded. Five short pauses were included in the experiment: one after the practice items and four during the experimental lists, resulting in one block of 16 practice items, four blocks of 80 experimental and filler stimuli, and a final block

of 85 experimental stimuli and fillers. The total duration of an experimental session was approximately 30 minutes.

*Results and discussion.*    Figure 1 shows the average percentages of errors (incorrect responses and time-outs) for the items correctly realised with voiceless final obstruents and for the items incorrectly realised with voiced final obstruents. The upper panel gives the scores for the plosive final items and the lower panel for the fricative final items. The percentages are broken down by the alternating character of the final obstruent.

Since the tokens ending in voiced obstruents were actually non-words, they elicited more "no" responses, and therefore more errors given our instructions, than the words ending in voiceless obstruents. For each item, we calculated the difference in the proportion of participants that produced
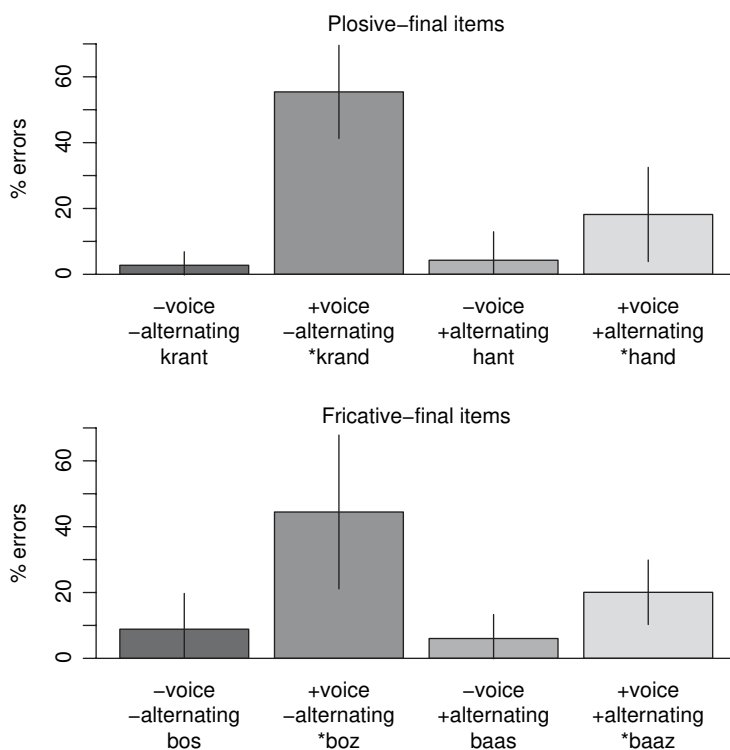


**Figure 1.**   Average percentages of errors (with standard error bars) for the stimuli ending in alternating and non-alternating plosives (upper panel) and fricatives (lower panel), realised as voiced or voiceless in Experiment 1.

TABLE 1
Mean differences in the error scores for words with alternating and non-alternating
final fricatives and plosives. Standard errors within parentheses

|  | Plosive | | | Fricative | | |
|---|---|---|---|---|---|---|
| Alternating | 0.139 | (0.169) | (*[hɑnd] – [hɑnt]) | 0.140 | (0.130) | (*[baz] – [bas]) |
| Non-alternating | 0.527 | (0.136) | (*[krɑnd] – [krɑnt]) | 0.356 | (0.192) | (*[bɔz] – [bɔs]) |

errors for the voiced and voiceless realisations. The means of these
differences are listed in Table 1.

We analysed the errors in two ways. First, we fitted a generalised mixed
effects model to the yes and no responses with participant and word as
crossed random effects (see below). This model contained a series of complex
interactions and for ease of exposition, we therefore report a simpler model
fit to the by-word difference scores. The two models yielded very similar
results.

The linear model for the by-word difference scores had as independent
variables Alternation (the alternating character of the word's final
obstruent: alternating versus non-alternating), Manner (the final obstru-
ent's manner of articulation: plosive versus fricative), and the word's log
CELEX Lemma Frequency (Baayen, Piepenbrock, & Gulikers, 1995). Since
the log lemma frequencies were not equally distributed over the words with
alternating (mean 7.04) and non-alternating (mean 6.07) words [$F(1, 92) =$
8.995, $p = .003$], we did not include interactions of Lemma Frequency by
Alternation. A step-wise model selection procedure (with removal of three
overly influential outliers with absolute standardised residuals greater than
2.0 and atypically high values for Cook's distance; Chatterjee, Hadi, &
Price, 2000, pp. 85–121) revealed main effects of Alternation [$\hat{\beta} = 0.187$,
$t(85) = 3.982$, $p < .001$, with $\hat{\beta}$ denoting the estimated (unstandardised)
regression coefficient, following the notation of Chatterjee et al.], Manner
[$\hat{\beta} = -0.605$, $t(85) = -4.269$, $p < .001$], and Lemma Frequency [$\hat{\beta} =$
$-0.038$, $t(85) = -2.323$, $p = .023$]. These main effects were modulated by
two 2-way interactions [Alternation by Manner: $\hat{\beta} = 0.267$, $t(85) = 4.455$,
$p < .001$; Manner by Lemma Frequency: $\hat{\beta} = 0.084$, $t(85) = 4.300$, $p < .001$].
This model explained 66% of the variance. Difference scores were greater
for words ending in non-alternating final obstruents, as predicted by the
hypothesis of paradigmatic effects in word recognition. Difference scores
were also greater for words ending in plosives than in fricatives, but only
for words ending in non-alternating obstruents, as expected given the weak
voiced–voiceless distinction of fricatives in Dutch.

Higher lemma frequencies led to greater difference scores for the plosive-
final words. Higher frequencies probably imply better knowledge of, or

faster access to, the canonical form of the words in isolation. By consequence, participants may have noticed the oddity of the final voiced plosive more quickly for higher frequency words, leading to an increase in the probability of a pseudoword response. Conversely, higher frequency led to reduced difference scores for the fricative final words. A voiced realisation may have been less odd for a final fricative because the opposition between voiced and voiceless realisations is almost completely neutralised in Dutch, and no longer morphologically distinctive. Apparently, a greater frequency renders a fricative final word more robust against variation in a non-distinctive feature.

The distribution of the reaction times, measured from word offset, was skewed with a number of very short outlier reaction times. We removed most of this skewness by applying a logarithmic transformation and by removing all reaction times (14, 0.5%) shorter than 50 ms (3.5 standard deviations below the mean). Figure 2 shows the average reaction times in milliseconds for the correct responses to the plosive final (upper panel) and the fricative final (lower panel) stimuli, broken down by the realisation of the final obstruent and its voice alternation. We analysed the log reaction times as a function of Realisation (the voice realisation of the final obstruent: voiced versus voiceless), Alternation (the obstruent's alternating character: non-alternating versus alternating), Manner (the obstruent's manner of articulation: plosive versus fricative), the Lemma Frequency of the word, and the rank of the stimulus in the experiment (its trial number), by means of a stepwise multi-level model of covariance (Baayen 2004; Baayen, Tweedie, & Schreuder, 2002; Bates & Sarkar, 2005; Pinheiro & Bates, 2000; Quené & Van den Bergh, 2004), with participant and word as crossed grouping factors. Multi-level modelling now offers the possibility of crossing multiple random effects and therefore obviates the necessity of separate $F_1$ and $F_2$ analyses. Note, furthermore, that by-item covariates, such as Lemma frequency, are confounded with the random effect of Word. As a consequence, our multi-level models are conservative with respect to the contribution of those covariates.

A step-wise model selection procedure revealed significant main effects of Realisation [$F(1, 2979) = 273.953$, $p < .001$, voiced words elicited longer reaction times], and Alternation [$F(1, 2979) = 6.837$, $p < .009$, alternating obstruents elicited shorter reaction times], which were modulated by an interaction of Realisation $\times$ Alternation [$F(1, 2979) = 30.178$, $p < .001$], and by an interaction of Realisation $\times$ Alternation $\times$ Manner [$F(2, 2979) = 13.150$, $p < .001$]. These interactions show that the difference between the voiced and voiceless realisations was significantly larger for the words ending in non-alternating obstruents, and especially so for non-alternating plosives. The greater inhibitory effect of a final voiced obstruent for words with non-alternating obstruents provides further evidence for
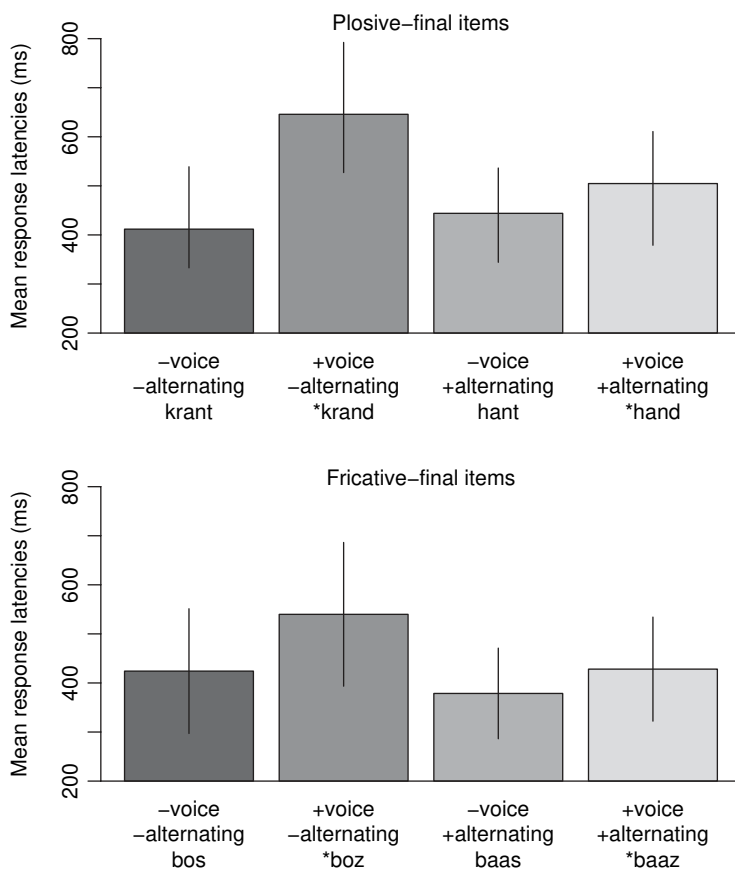
**Figure 2.**  Average reaction times (with standard error bars) for the correct responses to the stimuli ending in alternating and non-alternating plosives (upper panel) and fricatives (lower panel), realised as voiced or voiceless in Experiment 1.

paradigmatic effects in auditory word recognition. For words with alternating obstruents, a voiced realisation is supported by their morphological relatives in their inflectional and derivational paradigms, allowing faster responses.

With respect to the covariate, we observed a main effect of Lemma Frequency $[t(2979) = -16.862, \ p < .001$, words with a higher lemma frequency elicited shorter response latencies]. The random effects part of the multi-level model indicated that there were individual differences between the participants with respect to their sensitivity to Manner and Realisation (Log-likelihood ratio tests, $p < .001$). This is probably due to differences in linguistic background between the participants: Some participants may have

had a weaker voiced–voiceless opposition for fricatives than other partici-
pants. The $R^2$ of this model was 0.73.

The factorial contrasts in this experiment show that listeners produced
more errors, and that they reacted more slowly to realisations ending in
voiced plosives. Importantly, this was especially so for those words of which
the paradigms do not support the voiced realisation (the non-alternating
words). We conclude that both the accuracy measure and the response
latencies show paradigmatic effects.

To obtain a better insight into these paradigmatic effects, we investigated
whether the amount of paradigmatic support for a voiced realisation might
be a function of the frequencies of the forms with the voiced allophone and
the frequencies of the forms with the voiceless allophone. The more frequent
the forms with the voiced allophone relative to the forms with the voiceless
allophone, the greater the paradigmatic support could be for the voiced
realisation, and the faster the yes-responses in lexical decision. Similarly, a
high relative frequency of the voiceless allophone could speed up yes-
responses on forms correctly realised with the voiceless allophone.

We calculated the relative frequency of the voiced allophone for each
stem, conditioning on the frequencies of the members of both the inflectional
and the derivational paradigms. We refer to this conditional probability as
the paradigmatic likelihood of voicing (PLV). To give an example, for the
word oord "place", the PLV is the sum of the frequencies of the words
bejaardenoorden, kuuroorden, lustoorden, oorden, rustoorden, toevluchts-
oorden, vakantieoorden, woonoorden, which all end in the plural oorden
([o:rdə]), divided by the sum of the frequencies of the these same words plus
the frequencies of ballingsoord, bejaardenoord, genadeoord, herstellings-
oord, kuuroord, lustoord, oord, rustoord, toevluchtsoord, vakantieoord,
verbanningsoord, vluchtoord, woonoord, which all end in the singular oord
([o:rt]).

The PLV incorporates both inflectional and derivational competition
since the lexical evidence for the voiced or voiceless allophone is not
restricted to the inflectional paradigm of the base word itself, but extends to
the base word as it occurs across the lexicon. For independent evidence for
the integration of inflectional and derivational paradigms, see Moscoso del
Prado Martín, Kostić, & Baayen (2004).

The PLV ranged from 0.039 to 0.869 for the words ending in alternating
obstruents in the experiment. Note that the PLV is zero for words with non-
alternating obstruents. The correlation between PLV and Lemma frequency
is not significant ($r = -.197$, $p > .1$). We analysed the reaction times for the
words ending in alternating obstruents, including the PLV as a predictor.

In this new model, Realisation interacted with Lemma frequency [$F(2,
1598) = 90.438$, $p < .001$]: a higher Lemma frequency facilitated lexical
decisions, but only for words ending in voiceless obstruents. In addition,

the model showed interactions of Realisation $\times$ Manner $\times$ PLV, which had both a linear component [$F(4, 1598) = 10.754$, $p < .001$], and a quadratic component [$F(4, 1598) = 2.379$, $p = .050$]. As in the first analysis, participants differed in their sensitivity to Manner and Realisation (log-likelihood ratio tests, $p < .001$). In addition, Rank, that is, the position in the experimental list, differentially affected the response latencies to the different words (log-likelihood ratio test, $p = .034$).

The complex interaction of Realisation $\times$ Manner $\times$ PLV is visualised in Figure 3, which illustrates the partial effects of the PLV on the reaction times for the plosive final (solid lines) and fricative final (dashed lines) words, distinguishing between the voiced (triangles) and voiceless (circles) realisations, adjusted for the average lemma frequency (1125 per 42 million). We have an inverted U-shape curve for the voiced plosives, a flatter version for the voiced fricatives, a climbing curve for the voiceless plosives, and a descending curve for the voiceless fricatives.

The response latencies for the words ending in voiced obstruents are substantially longer than those to the words with voiceless obstruents, as in this case, we are actually dealing with yes-responses for pseudowords. At target offset, it was clear that the target was not a word, and the listener had to fall back on the residual activation in the cohort of lexical candidates. The most highly activated words in the cohort will have been those words which had the target as their initial constituent, irrespective of the voicing of the final obstruent. Henceforth, we refer to the set of words consisting of the target and its morphological continuation forms as the continuation set, to the subset of words with a voiced stem-final obstruent as the voiced continuations, and to the subset of words with a voiceless stem-final obstruent as the voiceless continuations. The continuation set of *[hɑnd] "hand" contains voiced continuations such as [hɑndə] "hands" and [hɑndəx] "handy", and voiceless continuations such as [hɑnt] "hand" and [hɑntsɑm] "manageable". Interestingly, the PLV is a measure of the competition between the voiced and voiceless continuations. If the PLV is approximately 0.5, competition is maximal, in which case the forms in the cohort are less activated, and hence less accessible, leading to delayed yes-responses. The closer the PLV is to zero or to one, the smaller the competition, and the faster the response latencies can be. The closer the PLV is to zero, the more a yes-response is based on the voiceless continuations. The closer it is to one, the more a yes-response is based on the voiced continuations. This explains the inverse U-shaped curves for the words with voiced final obstruents. The effect is attenuated for the words with voiced final fricatives. Given the weak voiced–voiceless opposition for fricatives, there is little competition between voiced and voiceless continuations, and the effects of PLV are smaller.
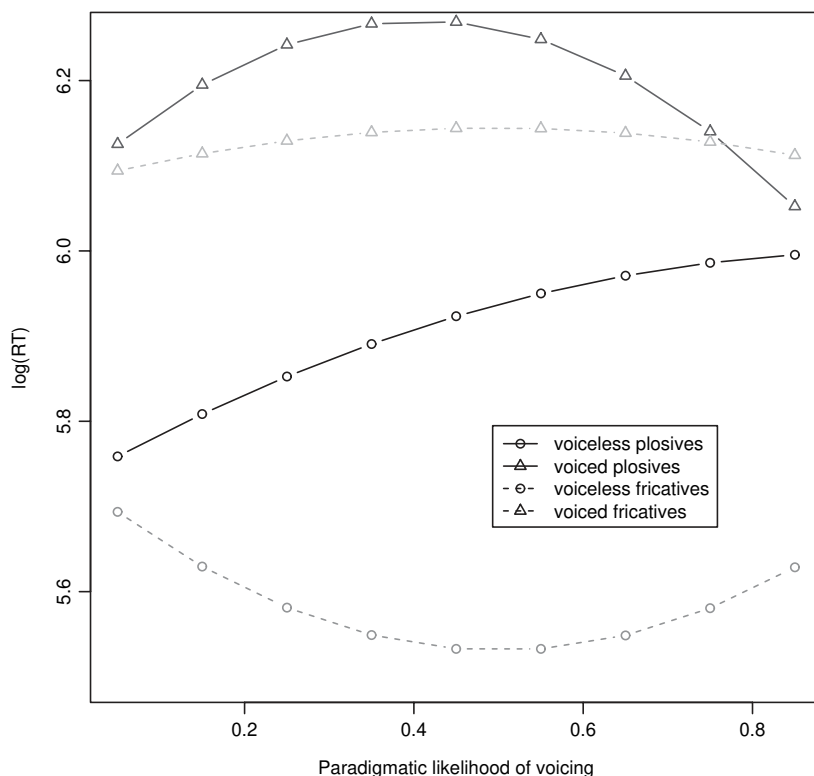
**Figure 3.** Partial effects of Realisation and Manner on log reaction time as a function of paradigmatic likelihood of voicing for the words with alternating obstruents in Experiment 1.

For the words ending in voiceless plosives, a response requires the deactivation of the voiced continuations. Conversely, the voiceless continuations support the final voiceless realisation. Again, the PLV is an estimate of the competition between the voiced and voiceless continuations. A higher PLV implies greater competition from the voiced continuations, and reduced support from the voiceless continuations, resulting in longer response latencies.

The pattern for the voiceless fricatives, probably requires a very different explanation. Here, given the weak status of the voiced fricative, the PLV seems to be not so much a measure of paradigmatic competition, but rather of paradigm size. A small PLV indicates that only one or just a few words, those ending in voiceless fricatives, are effectively present. The greater the paradigm size, the more lexical support and the shorter the response latencies. Note that the effect levels off quickly: as an indirect measure, the PLV quickly loses predictivity.

In conclusion, this experiment provides ample evidence for paradigmatic effects in word recognition. The data support the hypothesis that the recognition of a word, even a simple uninflected word, involves the activation of the word's complete morphological paradigm. The paradigmatic members compete among each other, and, depending on their similarity with the realisation of the target form, they lend support for its lexicality.

The acoustic measurements of the stimuli mentioned above (and reported in Ernestus & Baayen, in press b) show that the voiceless alternating plosives were weakly voiced. Weakly voiced obstruents are more similar to completely voiced obstruents than are completely voiceless obstruents, and as a consequence they may activate allomorphs ending in voiced obstruents to a greater extent. The effect of the PLV may thus be larger for realisations with weakly voiced obstruents than for realisations with completely voiceless obstruents. We tested this hypothesis in Experiment 2, a lexical decision experiment in which participants were presented with realisations ending in completely voiceless and weakly voiced final plosives.

## EXPERIMENT 2

### Method

*Participants.*    Forty native speakers of Dutch, most of them students at the Radboud University Nijmegen, were paid to participate in the experiment. None of them had participated in Experiment 1.

*Materials.*    We selected 97 monosyllabic words ending in alternating plosives from the CELEX lexical database. We tried to obtain realisations ending in weakly voiced and realisations ending in completely voiceless obstruents for these words as follows. We created a reading list for our speaker in which we presented each word twice, once written with a voiced final plosive, which is the correct spelling of the word (e.g., hand), and once with a voiceless plosive (hant). The correct spelling was preceded by a form of the word containing the allomorph with the voiced obstruent (handen), while the incorrect spelling was preceded by this same word form but written with a voiceless obstruent (hanten). We hoped that an incorrect spelling with the voiceless final obstruent (hant), in combination with the preceding phonologically related word form with the voiceless obstruent (hanten), would elicit completely voiceless realisations of the final obstruent, even though the word was highly similar to a word ending in an alternating obstruent (hand). In order to avoid list intonation on the second word of each pair, we had this actual target word followed by another word, such that each line contained three words. This third word was semantically related to the experimental word if this word was correctly spelled with a voiced

obstruent. The third word was semantically unrelated for words with a voiceless obstruent, in order to suggest that the experimental word was a pseudoword even though it was highly similar to an existing word. Thus our list contained lines like <u>handen hand pink</u> "hands hand little finger" and <u>hanten hant lamp</u> "pseudoword pseudoword lamp".

Acoustic measurements showed that our speaker had realised the final plosives that were spelled as voiced and those that were spelled as voiceless as approximately equally voiced. The two sets of experimental words hardly differed in the duration of their vowel [an average difference of 3 ms in the predicted direction, paired $t(96) = 1.486$, $p > .1$], or the duration of the release noise of the final plosive [an average difference of 2 ms in the opposite direction, paired $t(96) = 0.722$, $p > .1$]. Clearly our experimental manipulation failed. Nevertheless, the vowel duration and the duration of the release noise revealed a wide range of variation, which allowed us to select from the original 97 word pairs 23 pairs for which the acoustic measurements indicated a substantial difference in the amount of voicing between the members. These words are listed in the Appendix. We considered as weakly voiced those forms of which the vowel was at least 10 ms longer or the release noise at least 10 ms shorter than in the corresponding form in the pair. We labelled the corresponding forms as ending in completely voiceless final obstruents. The stimuli ending in weakly voiced plosives had longer vowels, on average 8 ms [paired $t(22) = 2.256$, $p = .034$], which is twice the difference in vowel length observed by Warner et al. (2004). In addition, our weakly voiced plosives also had shorter bursts, on average 19 ms [paired $t(22) = -4.411$, $p < .001$]. Compare the 23 ms difference observed by Ernestus and Baayen (in press a), and the 9 ms difference observed by Ernestus and Baayen (in press b). In this way, we obtained 23 pairs with a healthy difference in final voicing.

We created a first list containing one word from every pair: 11 words ending in weakly voiced obstruents and 12 words ending in completely voiceless obstruents. We created a second complementary list, which contained the remaining words. The words in each list were randomly mixed with 309 words that also functioned as filler words in Experiment 1. In addition, we added 30 filler words ending in voiceless non-alternating plosives. All words (experimental words and fillers) occupied the same position in the two lists. Both lists were preceded by 10 practice items.

*Procedure.*    The procedure was the same as in Experiment 1, except that the participants were not instructed on how to respond to realisations ending in completely voiced obstruents, as there were no such realisations in the experiment.

*Results and discussion.*    The error rate in this experiment was so low
(5.7%) that we did not further analyse the accuracy measure. We used the
same procedure as in Experiment 1 for removing most of the skewness in the
distribution of the response latencies, which were measured from stimulus
offset: a cut-off point at 50 ms (2.75 standard deviations below the mean)
removing 2.5% of the data points, and a logarithmic transformation. We
analysed the log reaction times for the correct responses by means of a
stepwise multi-level analysis of covariance with participant and word as
crossed grouping factors, and with as predictors the Voice Realisation of the
final obstruent (weakly voiced versus completely voiceless), the log lemma
frequency of the word, and the PLV. Inspection of the standardised residuals
led to the removal of six extreme outliers (data points with standardised
residuals larger than 3 and atypically high values for Cook's distance).

Words with completely voiceless plosives elicited shorter response
latencies [mean reaction time: 335 ms], than words with incompletely
neutralised plosives [mean reaction time: 365 ms; $F(1, 793) = 4.055$, $p =
.044$]. The faster processing of completely voiceless plosives may be due,
on the one hand, to the relative scarcity of incomplete neutralisation in
informal speech (e.g., Port & Crawford, 1989), and, on the other hand, to the
greater ambiguity of plosives with residual voicing. Lemma frequency did
not reach significance ($p > .1$).

The PLV emerged as a significant non-linear predictor [linear: $F(1, 793) =
8.806$, $p = .003$, quadratic: $F(1, 793) = 5.624$, $p = .018$]. This effect is
illustrated in Figure 4. The X-axis plots the PLV. The Y-axis shows the
predicted log reaction times adjusted for forms with weak voicing. The effect
of the PLV is similar to that observed for the stimuli with voiceless final
plosives in Experiment 1 (see Figure 3) in that responses were slower for
words with higher PLVs. A high PLV implies greater support for a voiced
realisation, and reduced support for the voiceless realisation of the final
plosive. As in Experiment 1, this slows listeners down. In the present
experiment, however, the effect of the PLV becomes visible only for words
with PLVs exceeding 0.5.

Recall that we designed this experiment to investigate whether the effect of
PLV observed in Experiment 1 might be modulated by the phonetic details of
the final obstruent. We hypothesised that the effects of the PLV might be
greater for more voiced final obstruents. The present experiment showed that
this is not the case. We observed an effect of incomplete neutralisation, and
effects of PLV, but no interaction between the two ($p > .5$). Apparently, the
effects of the paradigmatic neighbours arise independently of the fine
phonetic details of the voicing of the final obstruent in the stimulus. The
lexical paradigmatics emerge from this experiment as robust and as an
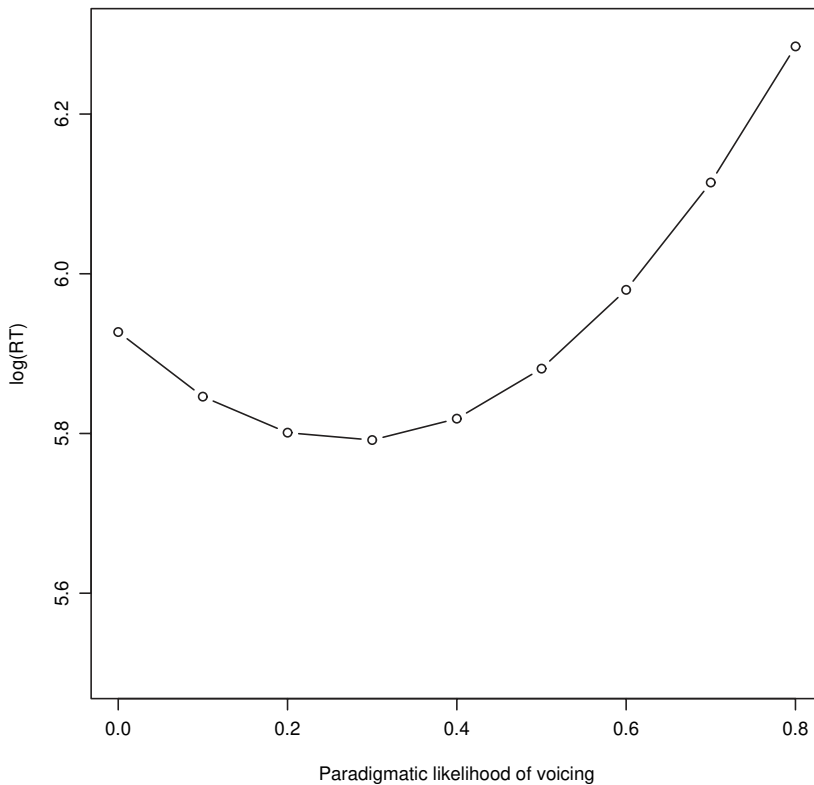inherent property of word recognition.

**Figure 4.** Partial effect of the paradigmatic likelihood of voicing on log reaction times in Experiment 2. The curve represents the words pronounced with weak voicing. For the words with completely voiceless final obstruents, the curve should be shifted by $-0.094$ units.

## GENERAL DISCUSSION

Many studies have addressed aspects of lexical competition between phonological neighbours in auditory word recognition (e.g., Goldinger et al., 1989; Luce, 1985; Pisoni et al., 1985). The present study addressed a specific kind of lexical competition that has received little attention so far, the competition between morphologically related words, that is, the competition in morphological paradigms. We focused on the recognition of Dutch morphemes ending in obstruents, which are obligatorily realised as voiceless in word-final position, but, depending on the paradigm, are voiced before word-internal vowels. That is, some morphological paradigms show a consistent voiceless realisation of the morpheme-final obstruent, with all members of the paradigm supporting the voiceless realisation (e.g., [krɑnt] with the plural [krɑntə]). In other paradigms, however, the final obstruent

has both voiced and voiceless allophones (e.g., [hɑnt] with the plural [hɑndə]), and the paradigmatic support for voicing (quantified in terms of the paradigmatic likelihood of voicing, PLV) varies with the frequencies of the different words in the paradigms. In two experiments, we investigated the processing consequences of the composition of the morphological paradigms with respect to the voicing characteristics of the stem-final obstruent.

In Experiment 1, word-final obstruents were correctly realised as voiceless, or incorrectly realised as voiced. Listeners were asked to perform lexical decisions and to ignore the voice realisations of the final obstruents. They performed better on incorrect realisations for those words of which the paradigms support the voiced realisation. Thus, they classified more often and also more quickly as words the incorrect form *[hɑnd] with the plural [hɑndə] than the incorrect form *[krɑnd] with the plural [krɑntə]. Both the responses and the response latencies thus show paradigmatic effects in word recognition: an incorrectly voiced realisation is recognised more easily if it is supported by other forms in the morphological paradigm.

This effect was modulated by the manner of articulation of the final obstruent. The voiced realisations were more problematic for the plosives compared to the fricatives. This came as no surprise, because the voiced–voiceless distinction is much stronger for plosives (e.g., Collins & Mees, 1981, p. 159; Gussenhoven & Bremmer, 1983, p. 57).

An analysis of the subset of words ending in alternating obstruents revealed a non-linear effect of the PLV. This effect varied with the voicing (correct or incorrect) of the final obstruent, as well as with its manner (plosive vs. fricative). For words with incorrect voiced plosives, the PLV maximally inhibited yes-responses when the likelihoods of the voiced and voiceless allophones were equal. This bears witness to the competition within the paradigm between the forms with voiced and voiceless obstruents. For words correctly realised with voiceless plosives, a higher PLV implied more inhibition, due to competition from the paradigm members with a voiced plosive.

In Experiment 1, participants heard voiced final plosives. These phono-tactically incorrect realisations may have affected the results for the correct realisations with voiceless plosives. Experiment 2 rules out this possibility. In this experiment, participants only heard words correctly ending in voiceless plosives. These plosives were either completely voiceless or slightly voiced due to incomplete neutralisation. Words with incomplete neutralisation elicited longer response latencies, just as the completely voiced obstruents in Experiment 1. Importantly, the effect of the PLV was inhibitory, as it was for the voiceless plosives in Experiment 1, and it showed no interaction with the voicing of the final obstruent.

Non-alternating voiceless obstruents are spelled with letters for voiceless phonemes and their spelling thus perfectly reflects their pronunciation.

Alternating obstruents, in contrast, are consistently spelled with letters for voiced phonemes. Thus, both hand and handen are spelled with d, even though hand is pronounced with [t] ([hɑnt]). It is becoming increasingly clear that orthographic representations are important in spoken word recognition (see e.g., Ziegler, Muneaux, & Grainger, 2003). The spelling difference between alternating and non-alternating obstruents may also help explain our finding that listeners recognise incorrect pronunciations with final voiced obstruents more accurately and faster for words ending in alternating obstruents. Orthographic representations, however, cannot be the source of the attested correlations between the PLV and response latencies for words ending in alternating obstruents, as all alternating obstruents are consistently spelled as voiced.

Thus far, we have interpreted paradigmatic effects as indicative of competition within the paradigm. An alternative interpretation at a more general level would be that the processing system is biased by the likelihood of voicing in a word's paradigm, without commitment to the idea that this bias necessarily is mediated by lexical competition.

The present findings have important consequences for theories of lexical representation and lexical processing. Within generative grammar, mor-pheme-final obstruents that are voiced before word-internal vowels are represented as (underlyingly) voiced. Thus, the singular [hɑnt] with the plural [hɑndə] would be represented as /hand/. Given the frequency effects observed for inflectional forms (see, e.g., Alegre & Gordon, 1999; Baayen et al., 1997, 2002, 2003; Bertram et al., 1999; Schreuder et al., 1999; Sereno & Jongman, 1997), it is unlikely that /hand/ would be the only lexical representative of the inflectional paradigm of hand. If both hand [hɑnt] and handen [hɑndə] are lexically stored, the final plosive of hand [hɑnt] needs not be stored as /d/, since the information that this plosive is voiced in the plural form is already present in the lexical representation of the plural form itself. Our finding that listeners perform better on [hɑnt] than on *[hɑnd] renders highly abstract lexical representations of the type /hand/ even more unlikely. Current generative accounts cannot explain why realisations that correspond perfectly to underlying representations, such as /hand/, supposedly stored in the mental lexicon, are so difficult to recognise. In addition, it is unclear how such abstract accounts might explain the relevance of the PLV, a measure based on the frequencies of the different forms in the word's paradigm.

The current results also resist explanation within the cohort model proposed by Lahiri and Marslen-Wilson (1991). This model accounts for the recognition of words with assimilated segments by assuming that assimila-tion involves the addition of phonetic features that are absent in the word's lexical representation. Thus, one might assume that the Dutch word hand has a lexical representation in which the final plosive is underspecified for

voice. This would account for the recognition of both [hɑnt] and the assimilated form [hɑnd], as in handboek "handbook". However, it does not account for the difference in recognition latencies that we observed for the surface forms [hɑnt] and *[hɑnd]. As pointed out by a number of studies addressing the comprehension of assimilated forms, assimilated segments are acceptable only in the contexts licensing the assimilation (e.g., Gaskell & Marslen-Wilson, 1996; Mitterer & Blomert, 2003). We conclude that the underspecification account of Lahiri and Marslen-Wilson is challenged by our data.

The paradigmatic effects observed in the experiments can readily be accounted for within models allowing separate lexical representations for the different (inflectional) word forms (see e.g., Blevins, 2003; Bybee, 2001). In such theories, *[hɑnd] activates the word forms /hɑnt/ and /hɑndə/, among others, while *[krɑnd] activates /krɑnt/ and /krɑntə/. Incorrect *[hɑnd] matches the inflectional word form /hɑndə/ better than incorrect *[krɑnd] matches /krɑntə/. The incorrect realisation *[hɑnd] thus receives larger lexical support, allowing faster yes-responses. The observed predictivity of the PLV is in line with probabilistic exemplar-based models (e.g., Goldinger, 1998; Pierrehumbert, 2001, 2003).

The effect of incomplete neutralisation observed in Experiment 2 is interesting in its own right. First, it fits well with the accumulating evidence concerning the relevance of fine phonetic detail for word recognition (Davis et al., 2002; Ernestus & Baayen, in press a; Kemps et al., 2005; Kemps, Wurm, Ernestus, Schreuder, & Baayen, 2005; Salverda et al., 2003; Spinelli, McQueen, & Cutler, 2003). Second, Warner et al. (2004) suggested that listeners would only take incomplete neutralisation into account if this would help them performing their task. The longer response latencies for realisations with incomplete neutralisation in Experiment 2 show that this is not the case: incomplete neutralisation provided no information about the lexicality of the stimuli (see also Ernestus & Baayen, in press a). Furthermore, the inhibition caused by weak voicing demonstrates that incomplete neutralisation, which appears to be most pronounced in careful speech (Port & Crawford, 1989), slows listeners. In other words, in those situations in which speakers wish to be maximally clear, they introduce subphonemic cues which makes their speech harder to process.

The consequences of morphological relations in the mental lexicon have thus far received little attention in the literature on auditory word recognition. Symptomatic of this state of affairs is the way in which cohorts were defined by Marslen-Wilson and Welsh (1978), who discarded morpho-logical continuation forms from the cohort counts a priori. Recent studies have begun to chart the effects of morphological relatives in auditory word processing. Kemps, Ernestus, Schreuder, and Baayen (2005) and Wurm et al. (in press) documented the role of morphological continuation forms after the

uniqueness point by means of entropy measures. Wurm et al. also reported the relevance of the morphological family size (Moscoso del Prado Martín et al., 2004; Schreuder & Baayen, 1997) for auditory word comprehension. What the present study adds to these studies is a demonstration of the importance of the phonological characteristics of the forms in a word's paradigm.

# REFERENCES

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, *40*, 41–61.

Baayen, R. H. (2004). Statistics in psycholinguistics: A critique of some current gold standards. *Mental Lexicon Working Papers,* 1–45.

Baayen, R. H., Burani, G., & Schreuder, R. (1997). Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. In G. E. Booij & J. v. Marie (Eds.), *Yearbook of morphology 1996* (pp. 13–34). Dordrecht: Kluwer Academic Publishers.

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

Baayen, R. H., McQueen, J., Dijkstra, T., & Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In R. H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 355–390). Berlin: Mouton de Gruyter.

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

Baayen, R. H., Schreuder, R., De Jong, N. H., & Krott, A. (2002). Dutch inflection: The rules that prove the exception. In S. Nooteboom, F. Weerman, & F. Wijnen (Eds.), *Storage and computation in the language faculty* (pp. 61–92). Dordrecht: Kluwer Academic Publishers.

Baayen, R. H., Tweedie, F. J., & Schreuder, R. (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.

Bates, D. M., & Sarkar, D. (2005). *The lme4 library.* http://lib.stat.cmu.edu/R/CRAN/.

Bertram, R., Laine, M., Baayen, R. H., Schreuder, R., & Hyönä, J. (1999). Affixal homonymy triggers full-form storage even with inflected words, even in a morphologically rich language. *Cognition*, *74*, B13–B25.

Blevins, J. P. (2003). Stems and paradigms. *Language*, *79*, 737–767.

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

Boudelaa, S., & Marslen-Wilson, W. D. (2004). Allomorphic variation in Arabic: Implications for lexical processing and representation. *Brain and Language*, *90*, 106–116.

Bybee, J. L. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Chatterjee, S., Hadi, A. S., & Price, B. (2000). *Regression analysis by example*. New York: John Wiley & Sons.

Charles-Luce, J. (1993). The effects of semantic context on voicing neutralization. *Phonetica*, *50*, 28–43.

Collins, B., & Mees, I. (1981). *The sounds of English and Dutch*. Leiden: Leiden University Press.

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.

Dinnsen, D. A., & Charles-Luce, J. (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics, 12*, 49–60.

Emmorey, K. D. (1989). Auditory morphological priming in the lexicon. *Language and Cognitive Processes, 4*, 73–92.

Ernestus, M., & Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language, 79*, 5–38.

Ernestus, M., & Baayen, R. H. (in press a). The functionality of incomplete neutralization in Dutch: The case of past-tense formation. In L. M. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Laboratory phonology 8*. Berlin: Mouton de Gruyter.

Ernestus, M., & Baayen, R. H. (in press b). Intraparadigmatic effects on the perception of voice. In E. J. Van der Torre & J. Van de Weijer (Eds.), *Voicing in Dutch*. Amsterdam: John Benjamins.

Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 144–158.

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

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbours of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28*, 501–518.

Gussenhoven, C., & Bremmer, R. H. (1983). Voiced fricatives in Dutch: Sources and present-day usage. *North-Western European Language Evolution, 2*, 55–71.

Kemps, R., Wurm, L. H., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes, 20*, 43–73.

Kemps, R., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005). Prosodic cues for morphological complexity: The case of Dutch noun plurals. *Memory and Cognition, 33*, 430–446.

Lahiri, A., & Marslen-Wilson, W. D. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38*, 245–294.

Luce, P. (1985). *Structural distinctions between high and low frequency words in auditory word recognition.* Ph.D. dissertation, Indiana University.

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

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

Meunier, F., & Segui, J. (2001). Cross-modal morphological priming in French. *Brain and Language, 81*, 89–102.

Mitterer, H., & Blomert, L. (2003). Copying with phonological assimilation in speech perception: Evidence for early compensation. *Perception and Psychophysics, 65*, 956–969.

Moscoso del Prado Martín, F, Kostić, A., & Baayen, R. H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition, 94*, 1–18.

Picrrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137–157). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Pierrehumbert, J. B. (2003). Probabilistic phonology: Discrimination and robustness. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probability theory in linguistics* (pp. 177–228). Cambridge, MA: The MIT Press.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS, Statistics and Computing*. New York: Springer.

Pisoni, D. B., Nusbaum, H. C, Luce, P. A., & Slowiaczek, L. M. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication, 4*, 75–95.

Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17, 257–282.

Port, R. F., & O'Dell, M. L. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455–471.

Quené, H., & Van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, 43, 103–121.

Reid, A. A., & Marslen-Wilson, W. D. (2003). Lexical representation of morphologically complex words: Evidence from Polish. In R. H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 287–336). Berlin: Mouton de Gruyter.

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.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, 37, 118–139.

Schreuder, R., De Jong, N. H., Krott, A., & Baayen, R. H. (1999). Rules and rote: Beyond the linguistic either-or fallacy. *Behavioral and Brain Sciences*, 22, 1038–1039.

Sereno, J., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory and Cognition*, 25, 425–437.

Slowiaczek, L., & Dinnsen, D. A. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, 13, 325–341.

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

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, 7, 263–272.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32, 251–276.

Wurm, L. H., Ernestus, M., Schreuder, R., & Baayen, R. H. (in press). Dynamics of the auditory comprehension of prefixed words: Cohort entropies and conditional root uniqueness points. *The Mental Lexicon*.

Ziegler, J. C, Muneaux, M., & Grainger, J. (2003). Neighbourhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language*, 48, 779–793.

# APPENDIX

*Experimental words for Experiment 1.*

1A. Experimental words ending in obstruents that are voiced in inflectionally related words:

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

1B. Experimental words ending in obstruents that are always voiceless:

<u>klap</u> "bang", <u>mep</u> "clout", <u>schep</u> "scoop", <u>stip</u> "dot", <u>strip</u> "strip", <u>beurt</u> "turn", <u>cent</u> "cent", <u>fluit</u> "flute", <u>geit</u> "goat", <u>grot</u> "cave", <u>hert</u> "deer", <u>kat</u> "cat", <u>klant</u> "customer", <u>knot</u> "knot",

krat "crate", kreet "cry", krot "slum", lat "slat", lint "ribbon", maat "measure", mot "moth", pet "cap", pit "pip", poort "gate", put "well", schat "treasure", scheut "twinge", spruit "sprout", staart "tail", bes "berry", bos "woods", dans "dance", eis "requirement", fles "bottle", kous "stocking", pols "wrist", tas "bag", tros "cluster", vis "fish", zeis "scythe", bef "jabot", juf "female teacher", nimf "nymph", plof "thud", rif "reef", slof "slipper", straf "punishment".

## Experimental words for Experiment 2.

band "band", blind "blind", bloed "blood", boord "border", bord "plate", brood "bread", bruid "bride", bult "lump", draad "thread", getob "brooding", glad "slippery", goot "gutter", groet "greeting", hoed "hat", huid "skin", jood "jew", kwaad "angry", kwab "lobe", miljard "thousand million", naad "seam", rib "rib", rond "round", schuld "debt", smaad "slander", spoed "speed", woud "forest".