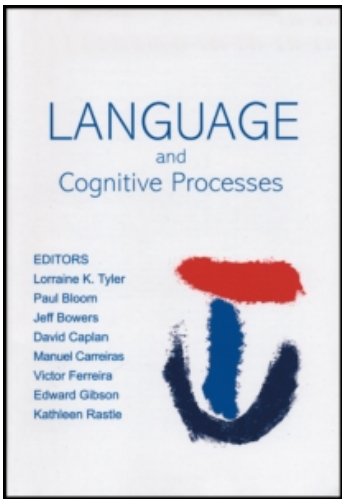

Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683153

Coping with gradient forms of /t/-deletion and lexical ambiguity in spoken word recognition

Esther Janse [a]; Sieb G. Nooteboom [a]; Hugo Quené [a]
[a] Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, the Netherlands

First Published on: 31 March 2006

To cite this Article Janse, Esther, Nooteboom, Sieb G. and Quené, Hugo(2006)'Coping with gradient forms of /t/-deletion and lexical ambiguity in spoken word recognition',Language and Cognitive Processes,22:2,161 — 200
To link to this Article: DOI: 10.1080/01690960500371024
URL: http://dx.doi.org/10.1080/01690960500371024

LANGUAGE AND COGNITIVE PROCESSES
2007, 22 (2), 161–200

Psychology Press
Taylor & Francis Group

Coping with gradient forms of /t/-deletion and lexical ambiguity in spoken word recognition

Esther Janse, Sieb G. Nooteboom, and Hugo Quené
Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, the Netherlands

This study investigates how listeners cope with gradient forms of deletion of word-final /t/ when recognising words in a phonological context that makes /t/-deletion viable. A corpus study confirmed a high incidence of /t/-deletion in an /st#b/ context in Dutch. A discrimination study showed that differences between released /t/, unreleased /t/ and fully deleted /t/ in this specific /st#b/ context were salient. Two on-line experiments were carried out to investigate whether lexical activation might be affected by this form variation. Even though unreleased and released variants were processed equally fast by listeners, a detailed analysis of the unreleased condition provided evidence for gradient activation. Activating a target ending in /t/ is slowest for the most reduced variant because phonological context has to be taken into account. Importantly, activation for a target with /t/ in the absence of cues for /t/ is reduced if there is a surface-matching lexical competitor.

INTRODUCTION

When listening to running speech, listeners continuously encounter assimilation, deletion and epenthesis phenomena, which cause the acoustic signal to deviate from the intended word's canonical form. Several studies have investigated how listeners relate these altered word forms to lexical representations in their mental lexicon, and what the nature is of these representations. This study investigates how listeners cope with deletion of word-final /t/ when they are recognising monomorphemic nouns embedded

Correspondence should be addressed to Esther Janse, Utrecht Institute of Linguistics OTS, Trans 10, 3512 JK Utrecht, the Netherlands. E-mail: Esther.Janse@let.uu.nl

We thank Heleen Hoekstra for her help with the CGN corpus, and Rob van Son for his help with the IFA corpus. Holger Mitterer and Mirjam Ernestus are thanked for their relevant contributions on /t/-deletion in Dutch.

http://www.psypress.com/lcp DOI: 10.1080/01690960500371024

in a phonetic context that makes /t/-deletion viable. The question addressed here is whether listeners actually perceive differences between changed and canonical forms, and whether the different variants are equally effective in activating the intended word. If differences between changed and unchanged forms are very difficult to hear, it seems unlikely that one variant would be less effective in activating the lexical item than the other. It makes sense that the greater the acoustic/perceptual difference between reduced and canonical forms, the greater the possible effect on immediate lexical activation.

Several studies on form variation and lexical access have investigated the process of place assimilation, particularly of nasals (as in *garden bench* in which /n/ surfaces as /m/). However, place of articulation of syllable-final nasals and plosives is notoriously difficult to hear, in particular in the context of a following consonant. Place cues of unreleased final stops are less salient than place cues in syllable onsets, and place cues are even weaker in the case of nasals. Thus, when unreleased stops or nasals occur before a stop, the weak place cues for the coda consonant tend to be perceptually masked by the more salient release of the following stop (Hura, Lindblom, & Diehl, 1992; Kohler, 1990; Ohala, 1990). Hence, this type of place assimilation might be a less suitable test ground to answer the question whether form variation slows down the word recognition process.

In the present study, therefore, a different connected-speech process was chosen. The process of /t/-deletion was investigated because deletion of /t/, be it partial or complete, was expected to be perceptually salient. Hence, this process might slow down lexical access. There is of course at least one important difference between assimilation and deletion, namely that assimilation may produce a phonemic mismatch between surface and underlying form, whereas deletion only yields a form with partial overlap. Based on partial overlap or relative goodness of fit, reduced forms will also activate the intended form, but to a lesser extent. It is possible that words are even recognised on the basis of the reduced input through lexical knowledge. It seems likely, however, that the phonological context following the target word can be used to enhance the initial amount of activation: listeners make use of the phonological regularity that /t/ may be reduced in this particular context. Hence, a word such as *feast* may be recognised in a phonological context licensing /t/ deletion, even in the absence of any cues for /t/. Recognition would then be context-dependent. How this recognition would work, however, is still an issue of debate.

One possibility is that recognition occurs via *compensation for deletion.* This compensation process has been argued to occur prelexically for place assimilation (Gaskell & Marslen-Wilson, 1998; Mitterer & Blomert, 2003). Even though there is an audible difference between the encountered speech segment and the 'underlying' segment, be it in terms of phonemic mismatch or in terms of missing phonemes, this difference is neutralised or

compensated for by the phonological context. Listeners may encounter, e.g., the sequence [mb], but the phonological context leads them to infer that /nb/ may have been intended. Mitterer and Ernestus (2006) performed a classification study with nonwords with varying degrees of /t/-lenition, embedded in different phonological contexts. Their results show that listeners, given weak cues for deleted /t/, report more /t/s in a nonword classification task in a context that licenses /t/-deletion than in contexts in which deletion of /t/ hardly ever occurs: listeners were more likely to infer the presence of a /t/ after /s/ than after /n/. Thus, phonological context is used to make up for a deleted /t/ segment. After prelexical reconstruction of the deleted or changed speech sound, a match can be made between a canonical template and the (reconstructed) speech signal. Mitterer and Ernestus (2006) claim that these results support the prelexical locus of compensation for /t/ deletion. On the basis of their crosslinguistic study, Gow and Im (2004) argue that coping with assimilation may require universal perceptual mechanisms rather than language-specific processes.

An alternative possibility is that recognition occurs because the reduced form is a stored exemplar of the intended word with final /t/ (Goldinger, 1998); accessing this form then is also licensed by a specific phonological context.

Either way, phonological context is taken into account in spoken-word recognition. This present study investigates lexical access. We will use the word detection task and repetition priming to tap the activation of lexical candidates. Therefore, we will not be able to draw any conclusions on the precise level of processing at which this context-sensitivity occurs. Anyhow, since taking into account post-target information takes time, the first hypothesis is that listeners will be slower in recognising words when presented with reduced input, in particular in full-deletion cases, than when they are presented with a more redundant input.

The second hypothesis is that listeners are less likely to activate targets with word-final /t/ in the absence of (strong) cues for /t/, if there is a surface-matching competitor (as in *chest*/*chess*) than if there is no such competitor (e.g., *feast*). If not, lexical ambiguity arises and listeners might recognise word forms that the speaker never intended. These two hypotheses are tested in the present paper.

In the generative Chomsky and Halle (1968) framework, speakers apply phonological processes in a categorical fashion: A phoneme is changed into another phoneme, or a phoneme is inserted or completely deleted. A number of speech production studies have shown, however, that many of these assimilation processes are better described as gradient. Browman and Goldstein (1992) describe assimilation processes as overlap between articulatory gestures in their Articulatory Phonology framework. In nasal place assimilation, there may be overlap between the tongue-tip gesture (necessary

for /n/) and the following lip-closing gesture (for e.g., /b/), resulting in forms in between /n/ and /m/, depending on the amount of overlap. With respect to /t/-deletion, there may be subtle acoustic cues that still hint at the 'deleted' segment. Browman and Goldstein (1990) provide X-ray evidence for 'gestural hiding', instead of categorical /t/-deletion. Two productions of the sequence *perfect memory* were analysed: one in which both words were read as members of a list, and one in which the sequence was part of a fluent phrase. In the fluent version, final /t/ of *perfect* was not audible, but the articulator movements suggested that the alveolar closure gesture for /t/ occurred in the fluent version, even with the same magnitude as in the word list version with a clearly audible final /t/. However, in the fluent version, the alveolar closure gesture was largely overlapped in time by other stop gestures: the closure portion by the preceding /k/; and the release portion by the following labial closure for /m/. Browman and Goldstein (1990) hypothesise that all casual speech alternations result from either an increase in gestural overlap, and/or a decrease in gesture magnitude.

This was confirmed by Nolan (1992), Gow (2001, 2002), and by Ellis and Hardcastle (2002) who show that nasal place assimilation often yields all types of forms that are intermediate between the canonical place of articulation and the assimilated place. Gow (2002) shows that the assimilated form [raɪ(t)p] taken from *right berries* is different from [raɪp] taken from the sequence *ripe berries,* even when only the most strongly assimilated items are selected. Gow's repetition priming experiment shows that underlyingly coronal primes ([raɪ^{t}p] taken from *right*) only primed coronal targets (*right*), but not noncoronal targets (*ripe*). Gow (2002) also shows that the assimilated signal in itself is ambiguous: assimilated coronals pattern with unmodified coronals in some acoustic measures, and with noncoronals in other measures, and are distinct from coronals and noncoronals in others. An additional repetition priming experiment in Gow's study (2002) shows that when the post-assimilation context is removed (which licenses the assimilation), the assimilated (underlyingly) coronal primes facilitate both coronal and non-coronal targets. This suggests that activation of targets that do not match the surface structure is enhanced by a viable phonetic context.

This is confirmed in a study by Mitterer and Ernestus (2006) on /t/-lenition in Dutch. Mitterer and Ernestus showed, by means of a nonword classification experiment that listeners take phonetic context into account when they interpret an ambiguous sound sequence. When presented with an item with lenited /t/ (e.g., *spes*[*t*] or *drin*[*t*]) and when forced to decide between two nonword alternatives (*spes*/*spest* or *drin*/*drint*), listeners reported /t/s more often in contexts where /t/ is often deleted. In other words, given the same lenited form of [t], listeners reconstruct /t/ more often after /s/ than after /n/, in accordance with the likelihood of deletion in production. Furthermore, following context also played a role: given an intermediate

form between *spes* and *spest,* listeners reported more /t/s when the following word started with /m/ than with /k/.

Different results have been reported pertaining to the question whether form variation affects processing speed. A number of studies have suggested that listeners recognise assimilated forms equally fast as canonical forms, as long as the assimilation is phonologically viable (Koster, 1987; Otake, Yoneyama, Cutler, & van der Lugt, 1996). Gaskell and Marslen-Wilson (1996) investigated how listeners cope with place assimilation (in nasals and plosives) in sequences such as *lean bacon* (pronounced /lim/). Their cross-modal repetition priming results showed equally large priming effects for the target word *lean* in the unassimilated condition (canonical *lean bacon* with /nb/) as in the assimilated condition (/mb/).

This is contradicted by the results of Coenen, Zwitserlood, and Bölte (2001), however, who investigated the perception of regressive and progressive place assimilation and progressive voice assimilation at word boundaries in German. Their results showed that although the assimilated forms activated the appropriate lexical representation, viably changed word forms always yielded smaller facilitation effects than unchanged word forms. In line with these results, Bard, Sotillo, Kelly, and Aylett (2001) showed that the presentation of phonologically reduced word forms, taken from spontaneous running speech, results in reduced activation of the intended word, as compared to the presentation of citation forms. Ernestus, Baayen, and Schreuder (2002) and Kemps, Ernestus, Schreuder, and Baayen (2004) have shown that highly reduced word forms (in their studies, taken from a spontaneous speech corpus) can be recognised by listeners only when they are presented in the context of several words (cf. also Pickett & Pollack, 1963). When they are presented in isolation or in a very limited context, listeners do not even recognise them. In line with this, more redundancy in word forms was shown to make fast speech easier to process (Janse, 2004). Given the same speech rate, neatly articulated (artificially time-compressed) speech can be processed faster by listeners than speech that was articulated fast by the speaker, even though the time-compressed speech is less natural.

These inconsistent results might be related to the fact that different connected-speech processes have been studied by the different researchers. Whereas Gow (2001, 2002) and Gaskell and Marslen-Wilson (1996, 2001) focused on place assimilation in syllable-final nasals and plosives, Coenen et al. (2001) studied voice assimilations as well. Bard et al. (2001), Ernestus et al. (2002), Kemps et al. (2004) and Janse (2004) looked at all sorts of reduction processes in spontaneous speech.

It makes sense that the larger the perceptual difference between the changed and canonical form, *given the specific phonetic context the word is embedded in,* the larger the effect on lexical access.

The results of Mitterer and Blomert (2003) are particularly relevant here because they provide brain response evidence that nasal place assimilation in a viable phonetic context is hardly picked up by listeners. In Mitterer and Blomert's brain response study, auditory stimuli were presented to listeners in a passive-listening ('odd-ball') task in order to obtain an event-related potential (ERP). Listeners are presented with a series of stimuli at a steady rate: there are only two different stimuli and one is much more frequent than the other. A negativity in the EEG signal then arises to the 'deviant' stimulus (e.g., assimilated *gardem bench*) as compared to the 'standard' stimulus (e.g., unassimilated *garden bench;* note that the study was performed in Dutch, but English examples are used here for the sake of readability). This mismatch negativity (MMN) is used as a measure of the extent to which the listener perceives a difference between the standard and the deviant (Näätänen & Winkler, 1999). Mitterer and Blomert showed that a MMN is only observed for the deviant form *gardem chair* (compared with standard *garden chair*), but not for deviant *gardem bench* (compared with *garden bench*). Since the acoustical differences are identical in the two comparisons, the MMN data provide evidence that auditory speech perception is context-sensitive. Mitterer and Blomert (2003) therefore argue that compensation for assimilation takes place at an early pre-lexical level and that general auditory processing mechanisms are responsible for perceptual integration, such that the sequence /mb/ is reanalysed as a possible /nb/ sequence (Mitterer, 2003).

Another option would be to argue that if changed and unchanged forms cannot be distinguished on an auditory level, not even at the early speech processing stage that Mitterer and Blomert (2003) investigated, the two forms are interchangeable and compensation is not necessary.[1]

Thus, the inconsistency in the literature whether form variation affects processing time may very well be related to the salience of the connected-speech process in question. The bigger the difference between changed and canonical forms, the more lexical access might be affected. In line with Coenen et al. (2001), Bard et al. (2001), and Janse (2004), the hypothesis is that more redundant phonetic variants of the word are processed faster than reduced forms because taking following phonological context into account is time-consuming.

[1] It is important to note that Mitterer and Blomert's results (2003) cannot fully explain the results obtained by Gaskell and Marslen-Wilson (2001) since the Gaskell and Marslen-Wilson study involved both place assimilation in nasals and plosives. On the basis of Mitterer and Blomert (2003), one might expect to find activation of RUN (and RUM), both in case listeners are presented with the sequence *run picks* and with *rum picks* (which should be homophonous). This is not what Gaskell and Marslen-Wilson found. However, they did not analyse the nasal cases separately from the coda plosive cases. It is conceivable that the acoustic differences between assimilated and unassimilated forms is greater for the nasals than for the plosives.

The present study tests these predictions by focusing on deletion of /t/ in one specific phonological context, namely in word-final /st/ clusters that are followed by a word-initial /b/. In this context, deletion of /t/ is said to occur frequently in English and Dutch, since both preceding /s/ and following /b/ reduce the probability that /t/ is released (for English: Guy, 1980; Neu, 1980; for Dutch: Booij, 1995; Ernestus, 2000; Mitterer & Ernestus, 2006; Zwaardemaker & Eijkman, 1928). If /t/-release is overlapped by the following /b/, there may be an audible trace of the closure for /t/. So, if /t/-deletion is gradient, subtle acoustic cues may still hint at underlying /t/ when no release is present. If these cues are picked up by listeners, lexical access may hardly be affected. Deelman and Connine (2001) found that presentation of word forms with word-final released stop consonants (as in *combat*) yielded equally high activation levels as forms with unreleased final stops. Their results therefore suggest that spoken word recognition does not require special mechanisms for processing unreleased variants. Sumner and Samuel (2005) also investigated variation in realisation of final /t/. They compared canonical (released) variants, variants with glottalisa-tion on the vowel and glottal coarticulation, and variants with a singly articulated glottal stop as the final phoneme. Their priming results showed that the strength of semantic priming was not influenced by phonetic variation in the prime: all regular variants equally and effectively primed the semantically related target.

However, if the deletion of /t/ is more or less complete in this specific /st#b/ context, listeners may still accept the word form as a variant of a word ending in /t/, given the phonological context. Activation is then initially based on partial overlap (cf. Deelman & Connine, 2001; Zwitserlood & Schriefers, 1995) and the following phonetic context information may consequently enhance the initial activation. Importantly, if listeners take this context information into account, they should run into trouble when the signal can be mapped onto multiple lexical forms: a lexical form with word-final /st/ (such as *chest*), and a form with word-final /s/ (*chess*). In other words, ambiguities may arise when listeners are confronted with nouns ending in /st/ that have a /s/ counterpart in the lexicon. Only top-down information is left then to decide which word must have been intended.

The question of lexical ambiguity has been addressed by Gaskell and Marslen-Wilson (2001) who investigated sequences such as *a quick run picks you up,* in which assimilated *run* yields *rum.* Whereas full 'compensation for place assimilation' was found in Gaskell and Marslen-Wilson's *leam bacon* study (1996), their (2001) results showed that the presence of a lexical item matching the surface form of speech blocks the contextual compensation process in a semantically neutral context. In other words, when listeners were presented with the *rum picks* sequence (surface form), priming was found only for the visual target *rum,* but not for *run.* However, when a biasing

context was provided, and given the appropriate phonological context for assimilation, a significant priming effect for *run* was found after the presentation of *rum picks.* Spinelli, McQueen, and Cutler (2003) also studied lexical disambiguitation by investigating how listeners process liaison cases in French (*dernier oignon,* in which /r/ surfaces, is almost homophonous with *dernier rognon*). Their results showed that there were subtle but reliable durational differences between the two versions of the lexically ambiguous phrases. Nevertheless, unintended words (*oignon* in *dernier rognon,* and *rognon* in *dernier oignon*) were 'weakly activated'. Thus, although the unintended candidate may not be completely blocked, the intended word is favoured.

So, given these results, the question is how salient /t/-deletion is in a viable deletion context, both if it concerns full deletion and if /t/ is unreleased. If differences between (partially) deleted and unchanged forms are picked up by listeners, processing time for the intended lexical item may be affected. The hypothesis is that full /t/-deletion delays spoken word recognition because following context is needed to license the reduced form as a variant of the target with /t/. A second hypothesis is that this licensing of a reduced form by context is decreased if a lexical competitor exists that matches the surface form.

The present study was set up to investigate first how often (noun-final) /t/-deletion occurs in everyday Dutch in this specific /st#b/ context. This issue was addressed by way of a corpus study of the Spoken Dutch Corpus (Oostdijk, 2000a, 2000b). If speech is tuned towards the listener, in line with the Hyper & Hypospeech theory (Lindblom, 1990), one might expect speakers to drop their noun-final /t/s less often if this would create potential ambiguities. Nolan (1992), however, showed that speakers also apply an assimilation process when a lexical alternative matches the assimilated form. Deletion rate was therefore studied both for nouns that have a lexical competitor without /t/, and for those that do not.

Secondly, a discrimination study was set up to investigate whether listeners are able to perceive the difference between forms with a fully released /t/, forms in which there is no /t/ at all, and intermediate forms. The greater the perceptual distance between reduced and canonical forms, the more likely it is that lexical access is influenced by the variation in form.

Thirdly, two on-line experiments were carried out to investigate whether form variation affects lexical activation. In order to study how listeners cope with (partial) /t/-deletion in their online processing of speech, a word monitoring experiment and a cross-modal repetition priming experiment were set up. 'Graded activation' was expected, in the sense that the more evidence is found for a certain target word, the stronger the activation of that word (Connine, Titone, Deelman, & Blasko, 1997; McQueen, Dahan, & Cutler, 2003). Hence, in terms of processing time, the most redundant form is

expected to yield the fastest processing times. With respect to lexical ambiguity, the hypothesis is that licensing of a reduced form by context is decreased if a lexical competitor exists that matches the surface form.

CORPUS STUDY

Two speech corpora were consulted to study how often /t/ is released in nouns with word-final /st/ in /st#b/ contexts. The first is the Spoken Dutch Corpus (CGN, release 5; Oostdijk, 2000a,b), the other is the much smaller speech corpus of the Amsterdam Phonetics institute (IFA corpus).

The Spoken Dutch Corpus is a speech database of contemporary Dutch as spoken by adults in Flanders and the Netherlands. The total number of words available in version 1.0 is nearly 9 million (800 hours of speech). Some 3.3 million words were collected in Flanders, and over 5.6 million words in the Netherlands. The corpus comprises a large number of samples of speech in several speech styles. The corpus was transcribed orthographically, lemmatised and enriched with part-of-speech information. For a small subset of the data, a broad phonetic transcription, a syntactic annotation, and a prosodic annotation were envisaged. Parts of the corpus have already been made available in the course of the project through a number of intermediate releases. In the present study, the intermediate release no. 5 was used.

Via an orthographic transcription search, all sequences of /st#b/ were looked up, where '#' stands for word boundary, in all speech styles (read, spontaneous, dialogue, etc.). Additionally, all sequences of word-final /st/ clusters followed by word-initial /m/ or /p/ were investigated. Only fragments of Dutch as spoken in the Netherlands were chosen.

Mitterer and Ernestus (in press) found that /t/-deletion, collapsed over different phonetic contexts, occurs more often in casual speech than in read speech. The present study focuses on /t/-deletion in /st#b/ contexts in monomorphemic nouns, which leaves only a restricted set of samples. This number of cases is too small to allow detailed analyses of speaker or style effects, and hence /t/-deletion cases will be counted only across several speakers and speaking styles. Presence or absence of the /t/ release constitutes the dependent variable; this was judged by ear and by visual inspection of the waveform and spectrogram. In some cases, a prosodic break occurred between the word-final /st/ and the following word with /b/. These breaks occurred most frequently in the broadcast news fragments and in the carefully read 'books on tape' fragments (read by volunteers for the visually impaired). Even though the insertion of breaks may be interesting in itself, since they may have been inserted in order to avoid ambiguity, this study was

set up to find out how often /t/ is released in running speech context. Therefore, the cases with a prosodic break are left out of consideration here.

Several categories of nouns were distinguished, namely nouns that were not ambiguous (did not have a /t/-less competitor, such as *feest* 'party'), nouns that were ambiguous (that did have a lexical competitor without final /t/, such as *kast*/*kas* 'cupboard/greenhouse'), and nouns that were proper names.

In total, 397 sequences were found via the orthographic search (172 with /b/, 21 with /p/, and 204 with /m/). Of 8 cases out of these 397 cases, the corresponding sound file was missing. Furthermore, in 99 cases, a prosodic break occurred between the noun ending in /st/ and the next word. This left 290 sequences in which there was no prosodic break between the noun-final /st/ cluster and the following word starting with a bilabial consonant. Of the 99 cases with a prosodic break, 89 had a release of /t/; in 10 cases /t/ was not released.

The results of the Spoken Dutch Corpus search for the 290 cases without prosodic break are presented in Table 1 below; these results are collapsed over /st/ clusters that are followed either by word-initial /b/, /p/, or /m/. The results are also collapsed over speakers and speech styles.

A X^2 test for equality of distributions was carried out. The distributions were not significantly different for the three categories ($X^2 = 0.51$, df $= 2$, $p > .1$). When only the ambiguous vs. nonambiguous nouns are entered into the X^2 test, there is no significant difference either ($X^2 < 1$, df $= 1$, $p > .1$).

If the search is restricted to /st#b/ sequences, 142 cases remain in which there is no prosodic break between the two words. Table 2 presents the data on whether /t/ is released or not in these sequences, broken down by noun type.

Again, collapsed over the noun categories, 15% of the cases contain a released /t/. When the results are broken down by category, the number of cases is fairly small, in particular for the ambiguous nouns. The distributions were not different for the three types of nouns ($X^2 = 3.30$, df $= 2$, $p > .1$).

TABLE 1
Number of cases in which word-final /t/ is either unreleased or released, in nouns with word-final /st/, occurring immediately before a word starting with /b,p,m/

	Unreleased	*Released*
Unambiguous nouns	201	33 (= 14% of total)
Ambiguous nouns	27	6 (= 18%)
Names (persons, places)	19	4 (= 17%)
Total	247	43 (= 15%)

TABLE 2
Number of cases in which /t/ is either unreleased or released, in nouns with word-final /st/, occurring immediately before a word starting with /b/ in the Spoken Dutch Corpus

	Unreleased	*Released*
Unambiguous nouns	101	14 (=12%)
Ambiguous nouns	9	3 (=25%)
Names (persons, places)	11	4 (=27%)
Total	121	21 (=15%)

Even in large corpora such as the Spoken Dutch Corpus, the number of nounfinal /st/ clusters preceding a bilabial consonant is too small for detailed analyses. The number of cases is too low to draw any reliable conclusions from it concerning the rate of /t/-release deletion for the ambiguous vs. unambiguous nouns.

The much smaller IFA corpus (van Son, Binnenpoorte, van den Heuvel, & Pols, 2001) comprises hand-segmented (Netherlands) Dutch speech from eight speakers in a variety of speaking styles (total of 50 000 words). An orthographic transcriptions search yielded 65 cases where a word-final /t/ was immediately followed by a word-initial /b/ or /m/. Since this corpus was much smaller, the search was not restricted to word-final /st/ clusters, but all types of word-final /t/s were examined. These were found in the speech styles 'spontaneous monologue', 'retell', and 'reading'. Collapsed across speakers and speaking styles, in 53 out of these 65 cases (82%), /t/ was not released.

The results of the two corpus studies clearly show that deletion of /t/ release in this specific context is widely attested in spoken Dutch (85% in the Spoken Dutch Corpus). The small number of cases does not allow a detailed analysis on the effect of speech style. Rate of /t/-deletion is probably lower in the more formal and more prepared speech styles (such as the broadcast messages). Still, the high percentage means that /t/ deletion in this phonetic context is widespread across the board. The limited number of cases also prevents an analysis on the effect of the presence of a lexical competitor.

Note that unreleased /t/ may still leave the listener sufficient cues to recover /t/. The absence of a /t/ release may not even lead to lexical ambiguity as long as other acoustic cues are present. Therefore, a discrimination experiment was set up to investigate whether differences between the forms with released /t/, unreleased /t/, and forms with no underlying /t/, can be perceived by listeners when these forms occur in a /st#b/ context. This experiment is described in the next section.

EXPERIMENT 1: DISCRIMINATION STUDY

Introduction

Mitterer and Blomert (2003) used a brain-response measure (Mismatch Negativity) to investigate whether listeners' brains register the difference between the assimilated form *gardem bench* and unassimilated *garden bench*. In a study on Hungarian liquid assimilation, Mitterer, Csépe, and Blomert (2003) used a (behavioural) discrimination paradigm to evaluate the discriminability of assimilated compared to unassimilated forms. The four-interval oddity (4I-oddity) task was used here because Gerrits (2001) and Schouten, Gerrits, and van Hessen (2003) show that this discrimination task shows the least influence of phonological categorisation on the stimuli (out of the most currently used discrimination tasks). Hence, listeners are equally likely to discriminate differences between phonological categories as differences within those categories, given the same acoustic difference. Mitterer's results (2003) showed that discrimination performance is context-sensitive: in the context in which the liquid assimilation occurred (viable context), deviations from a canonical form were less salient than in a nonviable context.

In the 4I-oddity task, which was also used here, for the same reasons, four stimuli are presented in a row at a constant interstimulus interval (ISI). Of these four stimuli, three are identical ('standard') and one is deviant ('odd'). The listener is told that the odd one is always the second or the third in the row of four stimuli. The listener is then asked to indicate which stimulus was odd: the second or the third one.

The central question in this part of the study was whether listeners are able to perceive the difference between forms with a fully released /t/, forms in which there is no /t/ at all, and intermediate forms when embedded in a context that makes deletion of /t/ viable. In order to decide on this issue, a discrimination experiment was set up in which listeners' attention was focused on differences between forms.

Method

Material. There were two categories of nouns in this study, namely the group of 'ambiguous' nouns that have a lexical competitor without word-final /t/, and the group of 'unambiguous' nouns that do not have such a lexical competitor. There are numerous nouns ending in /st/ in Dutch, but the number of nouns ending in /st/ having a competitor noun ending in /s/ is fairly limited. As a result, the ambiguous group contained only 12 nouns (e.g., *kast*/*kas* 'cupboard/greenhouse', *rust*/*Rus* '*rest*/*Russian*'); and the unambiguous group contained 24 nouns (e.g., *feest* 'party', *vest* 'cardigan').

With respect to the discrimination study, no differences were expected between the two groups of nouns, since the task is supposedly performed at a pre-phonological, and thus pre-lexical level (Schouten et al., 2003).

Semantically neutral sentences were constructed that could be used in the discrimination study and in the on-line listening experiment. Because this latter experiment was designed as a word-detection experiment, it is important that listeners should not be able to guess the crucial noun on the basis of the preceding context. For the ambiguous group of nouns, both lexical competitors should be possible continuations of the sentence (and ideally, they should be equally probable): cf. the example sentence in (1):

(1) *Hij kan beter de kast/kas bewaken*
('he can better the cupboard/greenhouse guard')

The target noun was always accented and was always followed by a verb starting with the unstressed initial syllable /bə/. This noun-verb sequence was chosen because it is highly unlikely that a speaker would insert a pause in between the noun and the verb, and because Dutch has a wide choice of verbs starting with the prefix /bə/.

All materials (nouns and sentences) are presented in the Appendix. The sentence materials were articulated by the first author (female native speaker of Dutch), in three different conditions, presented in (2) below. These three conditions were recorded for both the ambiguous and the unambiguous word items (note that the 'no /t/' condition might be artificial for the unambiguous words as it is uncertain whether a pronunciation without any trace of /t/ is realistic). In the ambiguous item set, the 'no /t/' condition is totally absent: this was intended by the speaker as a variant of the lexical word without /t/ (*kas*). The speaker tried to keep intonation and speech rate as similar as possible in the three versions to minimise confounding differences between them.

(2)	*unambiguous*	*ambiguous*
released /t/	vest beroven	kast bewaken
unreleased /t/	ves/t/ beroven	kas/t/ bewaken
no /t/	ves beroven	kas bewaken

There were no prosodic breaks between the crucial noun and the following verbs in either condition. In Figure 1 below, the waveforms of the three versions of the (unambiguous) target word *vest* (/vɛst/ 'cardigan') are shown, plus the first unstressed syllable of the following word *beroven* (/bə/ 'rob'). The three waveforms clearly show that, apart from the presence or absence of a release burst, there are temporal differences. The target word in the released version has the longest duration, but there is also a marked difference in closure duration between the *ves*/*t*/ and **ves* version.

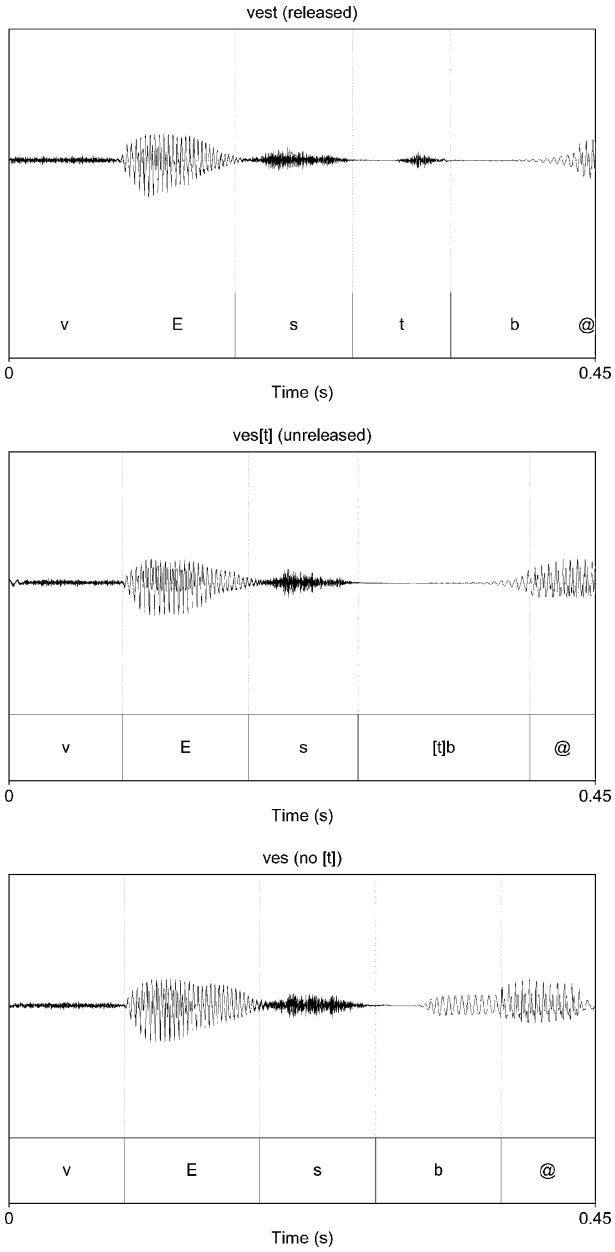

Figure 1. Waveform displays of target word *vest* and first syllable of next word *beroven* (/vɛstbə/, 'cardigan rob'), in three variants.

For the discrimination study, short fragments (containing the crucial noun plus the first syllable of the following verb; *vest be-*) were excised from the sentences (cf. Figure 1). If there was a slight difference between the forms that was not related to the crucial deletion of /t/, such as an audible difference in pitch during the target word vowel, cross-splicing of the vowel part was applied. Importantly, as can be seen in the waveform graphs above, the three variants of the word differed in duration, depending on the presence or absence of /t/. Furthermore, the acoustic characteristics of /b/ were related to the presence or absence of a /t/ release burst: weaker /b/ releases were found when the preceding /t/ was released than in case it was not released. Since these two aspects are inherent to the process of deletion, these differences were maintained for the sake of ecological validity of the variants.

Participants. Twelve listeners participated in the study, on a voluntary basis. They were students at Utrecht University and between 20–30 years of age. They reported no hearing difficulties.

Procedure. The three different acoustic forms of the word items were presented in the discrimination study as three pairs: (1) released vs. unreleased, (2) released vs. no /t/, (3) unreleased vs. no /t/. Each member of each comparison pair was presented both as the standard, and as the odd one in the row. This yielded two repetitions. The four speech stimuli per trial were presented with a 300 ms interstimulus interval, as presented in (3) below.

(3)

SOUND 1	SILENCE	SOUND 2	SILENCE	SOUND 3	SILENCE	SOUND 4
kast be-	300 ms	*kast be-*	300 ms	*kast be-*	300 ms	*kast be-*

The material was presented to subjects over headphones. Subjects were seated individually in a sound-treated booth. The listener was (correctly) informed that either the second or the third member of the train of stimuli is odd. The stimuli were presented in a computer-based environment: after the auditory presentation of the four stimuli, subjects had to click either of two buttons on the computer screen (labelled '2nd or '3rd'). There was a short practice session of four items before the actual test began. After the practice session, subjects could still ask questions if anything was unclear.

There was no time-out on the subject's button click response. Response times were also gathered (measured from the offset of the last speech fragment) because these might give some indication of how difficult the

decision was for the listener. Subjects were not given feedback on the correctness of their response.

Results

For the ambiguous items, there were 576 observations per comparison: 24 items; 2 repetitions per comparison; and 12 listeners. For the unambiguous items, there were 288 observations per comparison: 12 test items; 2 repetitions per comparison; and 12 listeners.

Table 3 presents the data of the 4I-oddity discrimination experiment. For both groups of items, the percentages correct are given, plus the mean decision time (1% truncated means are given), and the d′ value.

First of all, the results show that the percentages correct are high. The d′ data, and not the percentages correct, will be analysed since these d′ values are corrected for a possible bias subjects may have for either button (Kaplan, Macmillan, & Creelman, 1978; Macmillan & Creelman, 1991).

A Univariate ANOVA was conducted on the d′ data (there are no repeated measures because d′ is an aggregated measure), with Ambiguity and Comparison as fixed factors and subjects as random factor. Results show that the effect of Ambiguity is not significant, $F(1, 11) < 1$, *ns*, but that there is a main effect of Comparison, $F(2, 10) = 15.1$, $p < .001$. The interaction between Ambiguity and Comparison is not significant, $F(2, 10) = 1.1$, *ns*. This result was expected because the discrimination task is supposed to tap a pre-phonological level. A post-hoc Scheffé test showed that the d′ value for the largest acoustic difference (released vs. no /t/) was significantly higher than that for the released vs. unreleased comparison ($p < .001$), and significantly higher than that for the unreleased vs. no /t/ comparison ($p < .001$).

The response time data were also fed into ANOVAs, both with subjects and items as random factors, to investigate the effect of Ambiguity and Comparison. Since the experiment was self-paced (the next item would only

TABLE 3
Results of 4I-oddity experiment, broken down by item set and acoustic comparison

Item set	*Comparison*	*% correct*	*RT*	*d′*
Unambiguous items	released vs. unreleased	89	902	3.08
	released vs. no /t/	96	671	4.17
	unreleased vs. no /t/	88	951	2.75
Ambiguous items	released vs. unreleased	86	1025	2.60
	released vs. no /t/	97	701	4.69
	unreleased vs. no /t/	86	980	2.77

be presented once the subject had made a decision), some response times were fairly long because subjects were told that they could take a break. By only including the 99 percentile data in the ANOVAs, these very slow responses (exceeding the cut-off value of 4715 ms) were left out of consideration. Response times were transformed to 1/RT responses in order to make the data distribution more normal (Ulrich & Miller, 1994). Univariate ANOVAs were carried out on the 1/RT data. The effect of Ambiguity was significant by subjects only, $F_1(1, 11) = 6.16$, $p = .030$; $F_2(1, 34) = 2.57$, $p > .1$, but the effect of Comparison was highly significant, $F_1(2, 10) = 30.92$, $p < .001$; $F_2(2, 33) = 42.12$, $p < .001$. The interaction between Ambiguity and Comparison was not significant, $F_1(2, 10) < 1$, *ns*; $F_2(2, 33) < 1$, *ns*. Post-hoc analyses showed that responses were fastest in the released vs. no /t/ comparison (collapsed over ambiguous and unambiguous items 681 ms), compared with those in the released vs. unreleased comparison (943 ms; by subjects and items: $p < .001$), and compared with the unreleased vs. no /t/ comparison (960 ms; by subjects and items: $p < .001$). Responses in the released vs. unreleased comparison did not differ significantly from responses in the unreleased vs. no /t/ comparison (943 and 960 ms, respectively).

These results are consistent with the d′ data: comparison pairs with higher discriminability were accompanied by faster responses.

Discussion

Listeners were well able to discriminate the different variants embedded in this phonetic context: in particular when the acoustic difference is largest (released /t/ vs. no /t/). These results suggest that /t/-deletion is salient, contrary to previous results on nasal place assimilation (Mitterer & Blomert, 2003). Although there are no *behavioural* discrimination results available for nasal place assimilation, the Mitterer and Blomert MMN results (2003) show that discriminability of assimilated and unassimilated coda nasals preceding /b/ is relatively low. However, listeners *are* sensitive to the different acoustic forms in /t/-deletion. This leaves open the possibility that any informational cues contained in unreleased /t/ can be used immediately by listeners in on-line speech processing.

Given the salient difference between changed and unchanged word forms, the prediction was that intact forms are processed fastest. The lack of cues for /t/ is expected to slow down processing. This prediction was tested in a word detection study, described in the next section.

EXPERIMENT 2: WORD MONITORING STUDY

Introduction

By means of word detection or word monitoring (cf. Kilborn & Moss, 1996), the recognition process was tapped in an on-line way in order to investigate how quickly listeners map the incoming signal, which may be more or less 'canonical' in form, onto the intended word. The prediction was that the unreleased condition may be processed as fast as the released condition, but that the lack of any cues for /t/ slows down spoken word processing. Secondly, the word monitoring experiment was set up to investigate whether listeners also highly activate a target ending in /t/ in a /sb/ context if the speaker actually intended a lexical competitor without /t/.

Method

Material. The sentences from which fragments were presented in the discrimination study were used as listening material here. Each sentence was recorded in three different versions (released /t/, unreleased /t/, and no /t/). Obviously, in the released and unreleased conditions, a target word ending in /t/ was intended. For the unambiguous words, the target word with /t/ was also intended in the 'no /t/' condition (since there is no lexical alternative without /t/). However, note that for the ambiguous words, in the 'no /t/' variant, the word *without* /t/ was intended (*kas* was intended, not *kast*).

To avoid possible confounding differences between the three versions, the first part of the sentence, up to the onset of the target word, of one version was copied and spliced into the other two variants so that the three variants of each sentence were identical up to the target word.

Since there are two competing word candidates for the ambiguous words, there are six possible listening conditions for this set of items (three acoustic variants, times two detection targets, e.g., *kas* and *kast*). The six different conditions were balanced across the 12 ambiguous items over six experimental lists in a Latin Square design. The same acoustic variant was thus presented in two different conditions: once to a subject who was asked to monitor for *kas*, and once to a subject who was asked to monitor for *kast*. For the unambiguous words, there were only three different listening conditions, since the word detection target can only be a real word (**ves* is not a real word). These three conditions were also balanced across the same six experimental lists.

After the listeners had been presented with the total set of ambiguous and unambiguous items (mixed with filler items), the ambiguous items were presented again, this time in the same acoustic variant but now with the opposite target. If the repetition effect turned out to be too great, the datapoints of this second presentation could be ignored.

In the first part of the experiment, there were $12+24=36$ actual test items. In addition, several types of filler sentences were added: there were 17 close misses (sentences containing words that had initial overlap with the target, so that listeners would find out that they should not press the button after hearing the first two phonemes); 39 misses (target word did not appear in sentence); and 24 hit fillers (target words did appear in the sentence, but: (1) at a position different from the test items, or (2) targets were not nouns, or (3) targets did not end in /-st/).

In the second part of the experiment, in which the 12 ambiguous test sentences were repeated, 28 filler items were also repeated.

Participants. Sixty students of Utrecht University participated in this listening experiment: 10 listeners were randomly assigned to each of the six lists. They reported no hearing difficulties and had normal or corrected-to-normal vision, and were paid €5 for their participation.

Procedure. Listeners were instructed that they had to press a button as quickly and as accurately as possible, once they spotted the pre-assigned (visually presented) target word in the sentence that they were to hear over headphones. However, they were asked not to respond too quickly: for example, they might have to monitor for *prins* ('prince'), whereas the word in the sentence could then turn out to be *print* or *printer.*

Results

Monitoring experiments can involve several types of responses: correct hits, misses, and false alarms. In the present experiment, two types of 'false alarms' were possible: listeners responding to a word which turned out to be another word (embedded responses; listeners monitoring for *kas* responded during the presentation of *kast),* and 'false alarm' responses which could have been caused by compensation (listeners detected *kast* in the absence of cues for /t/). These types of 'errors' are indicated in the results table.

Detection times were measured from the onset of /s/ in the target word. Responses initiated before this point and response times below 100 ms from this point (3.5% of the responses) were excluded (both in the response time and detection rate analyses) because these responses must have been based on guessing: listeners cannot have taken the full acoustic information into account.

The onset of /s/ was chosen because word forms may deviate from this point onwards: duration of /s/ may vary and this may provide a subtle cue to whether /s/ is underlyingly word-final or not (Cutler & Butterfield, 1990; Klatt, 1976; Lehiste, 1972; Salverda, Dahan, & McQueen, 2003; Turk

& Shattuck-Hufnagel, 2000; cf. for /s/ in particular: Mitterer & Ernestus, 2006). Mean duration of /s/ was measured in the three acoustic forms.

In Table 4 the mean duration of /s/ is presented, broken down by item set and acoustic variant.

Table 4 shows that duration of /s/ in the 'no /t/' condition for the unambiguous items (103 ms) is longer than durations in the released and unreleased conditions (91 ms). This duration difference is much smaller for the ambiguous items. This is contrary to expectation. The results of a corpus study (Mitterer & Ernestus, 2006) show that deletion of /t/ does not give rise to a significant lengthening of preceding /s/ (mean lengthening in their study was +7 ms, which was not significant). If anything, duration differences were expected for the ambiguous items, since duration of /s/ could function as a subphonetic cue to disambiguate the two lexical competitors for the ambiguous items (note again that for the ambiguous words, in the 'no /t/' variant, the target *without* /t/ was intended).

Therefore, in the present sentence material, it seems that the speaker may have exaggerated the duration of /s/ in some 'no /t/' cases. The 'no /t/' variants of the unambiguous items are in fact unnatural: the speaker attempted to produce non-words and to avoid the suggestion of intermediate /t/s. This may have led to unnaturally long durations of /s/ in this acoustic variant in some cases, which renders them ecologically invalid as natural tokens of /t/ deletion. The mean duration of /s/ in the 'no /t/' condition is also longer for the unambiguous items than for the ambiguous items. A closer inspection of these durations showed that there were three specific items in which the duration of /s/ in the 'no /t/' condition was at least 1.2 times that in each of the other two conditions. These three items (one on each experimental list) were therefore excluded from further analysis (i.e., from behavioural results analyses). The mean duration of /s/ in the unambiguous item set is shown again in Table 5, after the exclusion of these items.

Table 5 shows that /s/ is still longest in the 'no /t/' condition, but note that, within the unambiguous item set, the 95% confidence interval around the

TABLE 4
Mean duration of /s/, broken down by item set and acoustic variant

Item set	*Acoustic variant*	*Mean duration*	*N*	*S.E. of mean*
Unambiguous items	released	92.1	24	2.42
	unreleased	90.5	24	2.11
	no /t/	103	24	1.65
Ambiguous items	released	92.8	12	3.25
	unreleased	87.3	12	2.64
	no /t/	94.4	12	3.39

TABLE 5
Mean duration of /s/ in the unambiguous item set (after exclusion of several items), broken down by acoustic variant

Acoustic variant	*Mean duration*	*N*	*S.E. of mean*
Released	92.4	21	2.75
Unreleased	91.1	21	2.35
no /t/	101	21	1.44

mean /s/ duration of the 'no /t/' variant almost overlaps with that of the released condition. The overall tendency for /s/ to be longer in the 'no /t/' condition may be the same weak lengthening effect as found in the lenited tokens in the Mitterer and Ernestus corpus study (2006).

As explained in the Method section, the ambiguous items were presented twice on each list: the second presentation involved the same acoustic variant as the first time, but now combined with the competitor visual word detection target. Because response times were significantly faster for the second presentation, and because almost all subjects told the experimenter that the repetition of sentences made clear to them that the target words could either have a /s/ or /st/ ending, the data of the second presentation were excluded from further analyses.[2]

In Table 6, the detection results are presented (mean detection time and detection rate), broken down by Target type (with or without /t/), item group (ambiguous vs. unambiguous items) and Acoustic Variant. If the detection data reflect false alarms, this is indicated in the table.

Early detection responses, given before the onset of /s/ or within 100 ms from /s/ onset, were excluded from this table and from the analyses (this removed 3.5% of the responses). Inclusion of these early responses would increase detection rates in all cells (apart from the two embedded false alarms cells) by 2–4% points.

The first hypothesis was that spoken word processing would be slowed down for reduced variants because taking following context into account to license the reduced variant takes time. The second hypothesis was that this licensing of a reduced variant by the phonological context is decreased if a lexical competitor exists that matches the surface form.

[2] Mean word detection time and detection rates were calculated for the first and second presentation. Detection rate decreased from the first to the second presentation [$N = 538$ (= 75%) for the first presentation; $N = 382$ (= 53%) the second time], and response time was slowed down [from 573 ms for the first presentation to 451 ms for the second presentation; ANOVAs on mean RT shows that effect of Presentation was highly significant: $F_1(1, 59) = 18.0$, $p < .001$; $F_2(1, 11) = 40.3$, $p < .001$].

TABLE 6
Word detection results: mean response time and detection rates (% of total) in parentheses, broken down by target type, item set, and acoustic variant

	Pre-assigned target has /t/	
Acoustic variant	*Unambiguous words*	*Ambiguous words*
1. released /t/	392 (96%)	464 (98%)
2. unreleased /t/	424 (96%)	446 (98%)
3. no /t/	486 (94%; compensation?)	492 (85%; compensation?)
	Pre-assigned target without /t/	
1. released /t/		273 (30%; embedded false alarms)
2. unreleased /t/		486 (41%; embedded false alarms)
3.no /t/		463 (90%)

Response time analyses will be provided first, followed by detection rate analyses.

Response times (measured from /s/ onset, and excluding response times below 100 ms) to the targets with /t/ (upper half Table 6) were subjected to an inverse transformation to make the data distribution more normal (Ulrich & Miller, 1994). Univariate ANOVAs were used, instead of Repeated Measures, because this latter type of ANOVA cannot deal with unequal numbers of observations (more unambiguous than ambiguous items). Experimental List was included as a between-subject factor in the F_1 analyses to account for the variability due to the specific set of combinations that subjects were presented with (Pollatsek & Well, 1995). The effect of Ambiguity was significant by subjects, $F_1(1, 54) = 13.40$, $p < .001$; $F_2(1, 31) = 2.82$, $p = .10$): listeners were somewhat slower in detecting ambiguous words than unambiguous words. This same effect was found by Andruski, Blumstein, and Burton, (1994) who investigated the effect of subphonetic voice onset time manipulations on the amount of lexical activation: the presence of a lexical competitor slowed reaction time latencies in all conditions.

The effect of Acoustic Variant was not significant, $F_1(2, 53) = 1.18$, *ns*; $F_2(2, 30) < 1$, *ns*. However, the interaction between Ambiguity and Acoustic Variant was significant, $F_1(2, 53) = 4.07$, $p = .019$; $F_2(2, 30) = 3.40$, $p = .034$. Because of this interaction, results were also analysed for the two item groups separately.

For the unambiguous item set (balanced over 3 experimental lists), the effect of Acoustic Variant was significant, $F_1(2, 56) = 5.52$, $p = .006$; $F_2(2, 19) = 14.29$, $p < .001$. Post-hoc comparisons showed that this was due to the significant difference between the released and 'no /t/' condition (by subjects

$p = .003$; by items $p = .009$). The other comparisons were not significantly different (all p values $> .1$). Thus, the first hypothesis is confirmed for these items: listeners were significantly slower in processing reduced forms than unchanged forms.

For the ambiguous item set (targets with /t/; design balanced over 6 experimental lists), the effect of Acoustic Variant was significant by subjects only, $F_1(2, 53) = 3.45$, $p = .039$; $F_2(2, 10) = 2.39$, $p > .1$. The differences between the variants are obviously smaller than in the unambiguous item set.

Thus, word detection times were not slower in the unreleased condition: listeners are equally fast in processing released and unreleased variants. Full deletion, on the other hand, did slow down the recognition process, but mainly for the unambiguous items.

The hypothesis concerning lexical ambiguity will now be investigated by comparing the results for the two detection targets (upper and lower right quadrants in Table 6). Response times were only analysed for the 'no /t/' subset of the data, because the response time means for targets without /t/ in the released and unreleased conditions are based on false alarms. Given the same auditory presentation ('no /t/', intended as the word without /t/), response times were equally fast for targets without /t/ (463 ms) as for targets with /t/ (492 ms); the effect of Target is not significant in ANOVAs on the inverse response times in the 'no /t/' condition data subset, $F_1(1, 59) = 1.74$, *ns*; $F_2(1, 11) < 1$, *ns*. A t-test showed that detection rate for the target with /t/ (85%) was equally high as for the target without /t/ (90%; $t_1(59) = 1.10$, *ns*; $t_2(11) < 1$, *ns*.

A general detection rate analysis was performed to test the hypothesis that, given the 'no /t/' variant, licensing of a reduced form by the phonological context is decreased if a lexical competitor exists that matches the surface form. Detection rates for targets with /t/ were subjected to subject and item ANOVAs to study the effects of Ambiguity (a nesting factor in the item analysis) and Acoustic Variant. There was no significant effect of Ambiguity, $F_1(1, 59) = 1.04$, *ns*; $F_2(1, 31) < 1$, *ns*, but the effect of Acoustic Variant was significant, $F_1(2, 58) = 7.01$, $p = .002$; $F_2(2, 30) = 6.38$, $p = .005$. Most importantly, the interaction between Ambiguity and Acoustic Variant was significant in both analyses, $F1(2, 58) = 4.81$, $p = .012$; $F_2(2, 30) = 3.59$, $p = .040$. Clearly, detection rates were lower in the 'no /t/' condition, particularly for the ambiguous items.

A detection rate analysis (after arcsine transformation of rates) was performed on the ambiguous data set to test the effects of Target and Acoustic Variant. Listeners reported that, when they were monitoring for targets without /t/, they would sometimes press the button before they noticed that the word had not finished yet. Because then the target word is in fact embedded in the longer word, false alarm rates mainly reflect too hasty responses (30% in released condition, 41% in unreleased condition). The

ANOVA results showed significant main effects both of Target, $F_1(1, 59) = 209.7$, $p < .001$; $F_2(1, 11) = 74.7$, $p < .001$, and of Acoustic Variant, $F_1(2, 58) = 24.2$, $p < .001$; $F_2(2, 10) = 11.6$, $p = .002$. Clearly, the interaction between Target and Acoustic Variant was also significant, $F_1(2, 58) = 76.3$, $p < .001$; $F_2(2, 10) = 65.9$, $p < .001$. For the targets without /t/, more false alarms were found in the acoustically more ambiguous unreleased condition (41%) than in the released condition (30%), but a pairwise comparison showed that this difference was not significant, $t_1(59) = 1.90$, $p = .063$; $t_2(11) = 1.61$, $p > .1$.

Taken together, these results show that released forms and unreleased variants are equally effective in activating word forms with /t/. Recognising words ending in /t/ on the basis of a fully /t/-deleted variant required extra processing time, however, but mainly for the unambiguous items. Nevertheless, detection rate analyses provided evidence that the 'no /t/' condition is less effective in activating targets ending in /t/ for both item sets. Furthermore, the detection rate results showed that activation of targets with /t/, given a 'no /t/' variant, was affected by the presence of a lexical competitor: detection rates were mainly compromised if there is a lexical competitor. Thus, the only indication that this reduced form is less effective in activating the target with /t/ for the ambiguous items is in the detection rates. Still, given the 'no /t/' variant, detection time and detection rate did not differ for the two possible word targets.

Discussion

The results of the word monitoring study have shown that responses were not slower in the unreleased condition: listeners are equally fast in processing released and unreleased variants. This agrees with other lexical activation results elicited by different /t/ variants (cf. Deelman & Connine, 2001; Sumner & Samuel, 2005). However, word forms with full /t/ are processed faster than forms without cues for /t/. This confirms the hypothesis that form variation affects processing time. Listeners either need more time, or the licensing phonological context, to initiate their response in the absence of cues for /t/ (cf. Zwitserlood & Schriefers, 1995 on how extra processing time and additional segmental information are often confounded in auditory word recognition).

However, the detection results were not very clear with respect to the second hypothesis on the activation of targets with /t/, given a 'no /t/' variant, when there is a surface-matching competitor. In their lexical ambiguity study, Gaskell and Marslen-Wilson (2001) found that compensation for place assimilation was blocked if there was a lexical item corresponding to the surface form. However, in the present word detection data, when listeners were presented with a 'no /t/' condition in which a lexical candidate *without*

/t/ was actually intended, detection rates both for targets with /t/ and for targets without /t/ were high, but not 100%. Even though the present detection rate results might reveal a slight (insignificant) bias towards the surface-matching form (rate is 90% for target without /t/ and 85% for the target with /t/, given the same acoustic 'no /t/' variant), there is no clear evidence, however, that either candidate is favoured.

Two possible accounts for these 'equal activation' results can be given. The first account is that these results are an artifact of the word detection task. An obvious disadvantage of the word detection task is that the to-be-recognised word is given away before it is presented auditorily. This primes the recognition of this pre-assigned word. Combined with an emphasis on fast response times, and listeners' eagerness to respond (rather than not respond) in a go-no go task, this priming effect might blur differences between conditions, if these exist. Furthermore, false alarm rates in the 'embedded' conditions (lower part of Table 6) were as high as 41%. This raises questions about the validity of the results.

The second account is that these results show that the target with word-final /t/ is activated, even if there is a lexical candidate matching the surface form. Obviously, based on onset overlap, unintended *kast* should be active as well as intended *kas* up to some point during /s/. In fact, all detection responses that are initiated up to that point are legitimate. The question is whether the following phonological context still licenses activation of the target with /t/ to the same degree as when there is no such competitor. The present results suggest that listeners reconstruct a word form that the speaker never intended.

In order to be able to choose between these two accounts for the data, a different experimental paradigm was chosen in which the auditory stimulus is not primed beforehand. Gaskell and Marslen-Wilson (1996, 2001), Gow (2002) and Coenen et al. (2001) used cross-modal repetition priming to investigate lexical activation following the auditory presentation of changed word forms. For the sake of comparison, it seemed appropriate to set up a cross-modal repetition priming experiment with the same material set. In the word detection study, the pre-assignment of a word target may have influenced the spoken word recognition process. By presenting a certain visual target before auditory presentation, one particular lexical candidate is biased over other candidates. In a repetition priming set-up, however, there is no pre-assignment of targets. Visual targets are shown only after the auditory presentation. In this way, the auditory presentation of the different acoustic variants leads to certain activation levels for word candidates. These activation levels are then tapped by response times to the visual targets.

EXPERIMENT 3: CROSS-MODAL REPETITION PRIMING STUDY

Introduction

In the previous experiment, two hypotheses were tested. The first hypothesis was that more redundant forms can be processed faster than reduced forms. This hypothesis was confirmed by the detection time results, but only for full deletion and only for the unambiguous items. The detection rate analysis provided further support for this first hypothesis (lower rates in the 'no /t/' condition for both item sets). The second hypothesis was that licensing of a reduced form by the phonological context is reduced if a lexical competitor exists that matches the surface form. Because the word detection results did not provide clear evidence with respect to this second hypothesis, a cross-modal priming experiment was set up to investigate lexical activation once again. Even though a possible task artifact may have influenced the results of the detection study, this task artifact cannot explain the significant reaction time difference between reduced and more redundant forms. Therefore, the present experiment was mainly set up to investigate lexical ambiguity. Nevertheless, the same acoustic conditions were tested to investigate lexical activation elicited by the different acoustic variants.

One marked difference between word detection experiments and a cross-modal design is that a cross-modal design entails a particular choice with respect to the point at which the visual target is presented, relative to the auditory presentation. Given that the present study involves /t/-deletion in a particular phonetic context, activation of word candidates was studied following this licensing context, i.e., after /b/. The word detection results showed that detection times were slowest in the conditions in which the prime word itself was shortest (cf. Figure 1): listeners may have needed the following context to license this variant as an exemplar of a target with /t/. Even when target presentation is time-locked to the offset of /b/, activation levels for the intended word may differ for the acoustic variants. This was tested in Experiment 3. Secondly, the hypothesis was tested that, given the 'no /t/' condition, activation for the target with /t/ is decreased if the lexical item has a surface-matching competitor.

Method

Materials. The sentences used in the word detection study were also used as listening material here. However, additional baseline sentences were required in which the auditory sentence contained a control word instead of a target word. The repetition priming effect can then be evaluated by comparing the lexical decision times to the visual targets when subjects were presented with either of the test conditions with those in the baseline

conditions. The control words were monosyllabic words which were matched to their corresponding test word (or test word pair for the ambiguous items) in terms of frequency and cloze probability. Phonemic content of each control word did not overlap with that of its corresponding test word and there was no semantic relation between control word and visual target. Some control words ($N = 11$ out of 36) were CVC words; the other control words were either CCVC ($N = 7$) or CVCC ($N = 13$) or CCVCC ($N = 5$).

To avoid possible confounding differences between the material and the control sentences, care was taken that the control sentences were highly similar in style, rate, and intonation to the test material. The resulting recording was similar enough to allow cross-splicing without audible disruptions of the sentences: the first part of the sentence, up to the onset of the target word, of one version was copied and spliced into the control variant. The same cross-splicing procedure was done with the part of the sentence following the target/control word. The filler sentences were also rerecorded in the 'control' recording. In order to check that the 'test' and 'control' recording would not differ with respect to the visual target response times, six filler sentences were presented to half of the subjects in the 'test recording' version, and the other 30 subjects were presented with the 'control recording' version of these six sentences, in combination with the same visual target presented at the same point in the sentence. Response times (following a 1/RT transformation) to the same visual targets did not differ significantly between subjects presented with the 'test' recording versus those subjects presented with the 'control' recording ($p > .1$).

Because there are two word candidates for the ambiguous words, there were now eight possible listening conditions for this set of items (4 acoustic variants, i.e., 3 test variants plus one control condition, $\times 2$ detection targets, e.g., *kas* and *kast)*. Since eight conditions cannot be balanced across 12 ambiguous item pairs, two combinations were left out of consideration. The word detection results clearly showed that hit rates for targets without /t/ were relatively low when listeners were presented with conditions with released or unreleased /t/. In other words, positive acoustic cues for /t/ clearly (and immediately) rule out the possibility that the word to be recognised has no /t/ (and note that visual targets are only shown following the initial /b/ of the next verb). Therefore, only the control condition and the 'no /t/' condition were presented in combination with targets without /t/.

The six remaining conditions were balanced across the 12 ambiguous items. Combining six conditions (for the ambiguous items) and four conditions (for the unambiguous items; since there is only one visual target) in one experiment could only be achieved by using 12 experimental lists. The items and conditions were balanced across these lists according to a Latin Square design.

There were 12 (ambiguous items) + 24 (unambiguous items) = 36 actual test items on each list. In addition, there were 70 filler sentences to avoid strategic processing. Position of the targets in relation to the auditory sentence was varied to avoid anticipation. The 36 test + 70 filler sentences together were balanced for lexical status: 53 real words eliciting a YES response vs. 53 non-words eliciting a NO response. Test and filler sentences were also balanced for overlap with the target (in 53 sentences there was at least partial overlap with the target, in 53 there was no phonological overlap at all). Thus, subjects were confronted with all possible combinations: real word target with overlap; non-word target with overlap; real word target without overlap; and non-word target without overlap.

Participants. Sixty students of Utrecht University participated in this listening experiment: five listeners were randomly assigned to each of the 12 lists. They reported no hearing difficulties and had normal or corrected-to-normal vision, and were paid €5 for their participation. None of them had participated in the word monitoring experiment.

Procedure. Listeners were presented with the auditory material over headphones, while seated in a sound-insulated booth. They were instructed to listen to the auditory material and to make a lexical decision as quickly and as accurately as possible, once they were presented with a visual target on the computer screen in front of them. Visual targets were presented at the onset of the schwa in the following verb /vɛst↓ə/. Targets remained visible for 200 ms. Lexical decision times were measured from the onset of the target's visual presentation.

Results

The results are presented in Table 7, for both item sets and for both visual targets. The results concerning the targets with /t/ (upper part Table 7) will be discussed first.

The hypothesis that form variation affects lexical activation mainly concerns the upper part of Table 7. Response times were transformed to inverse RTs in order to make the data distribution more normal. ANOVAs were carried out on the inverse RTs (only including correct responses) to test the effects of Ambiguity and Acoustic Variant for targets with /t/. Subjects were nested under the factor Experimental List in the subject analyses. The effect of Ambiguity was significant: responses were slower in the (primed) ambiguous conditions, $F_1(1, 59) = 16.63$, $p < .001$; $F_2(1, 34) = 6.77$, $p = .013$. Thus, one can see the effects of lexical competition on activation levels even after the release of /b/. The effect of Acoustic Variant was highly significant as well, $F_1(3, 57) = 21.38$, $p < .001$; $F_2(3, 32) = 16.80$, $p < .001$, but there was

TABLE 7
Mean lexical decision times in ms (for correct decisions only) and error percentages, broken down by Target type, Item set, and Acoustic variant. Test-control difference RT scores are given in parentheses

	Target with /t/			
	Unambiguous items (e.g., vest)		*Ambiguous items (e.g., kast)*	
Acoustic variant	*RT (ms)*	*Error (%)*	*RT (ms)*	*Error (%)*
control sentence	676	5.7	684	5.8
released /t/	590 (86)	1.7	637 (48)	0
unreleased /t/	595 (81)	1.4	630 (54)	2.5
no /t/	616 (60)	0.8	695 (−11)	3.3
	Target without /t/			
	Ambiguous items (e.g., kas)			
control sentence			743	5.0
no /t/			674 (69)	9.2

no significant Ambiguity by Acoustic Variant interaction, $F_1(3, 57) = 1.83$, $p > .1$; $F_2(3, 32) = 1.17$, $p > .1$.

ANOVAs on the effect of Acoustic Variant and planned comparisons were carried out for the two item groups separately. For the unambiguous items, the effect of Acoustic Variant was highly significant, $F_1(3, 54) = 28.86$, $p < .001$; $F_2(3, 21) = 28.34$, $p < .001$. Planned comparisons (Scheffé) were carried out to check whether each test condition differed significantly from the control condition. The repetition priming effect was significant in all three test conditions (by subjects and items, all p values $< .001$).

An important issue here is the effect size: does the released acoustic version activate the intended word to a stronger degree than the 'no /t/' version? In order to compare priming effect size, difference scores were calculated by subtracting mean 1/RT in either of the test conditions from that in the control condition for each subject and for each item. T-tests were performed on these difference scores. The priming effect in the released condition (86 ms) was numerically larger than that in the 'no /t/' condition (60 ms), but this difference failed to reach significance, $t_1(1, 59) = 1.79$, $p = .079$; $t_2(1, 24) = 1.38$, $p = .18$.

For the ambiguous items, response times to the targets with /t/ were analysed in a similar way. The effect of Acoustic Variant was significant, $F_1(3, 52) = 7.03$, $p < .001$; $F_2(3, 9) = 10.78$, $p = .002$. Again, planned comparisons were performed to see which test conditions differed significantly from

the control condition. Response times in the released condition differed significantly from the control condition ($p < .001$ by subjects; $p = .019$ by items). Response times in the unreleased condition were also significantly different from the control condition ($p = .002$ by subjects; $p = .037$ by items). Crucially, there was no significant priming in the 'no /t/' condition: response times in the 'no /t/' condition did not differ significantly from the control condition ($p > .1$ by subjects and items). Lastly, the priming effect in the released condition (48 ms) was equally large as that in the unreleased condition (54 ms): *t*-tests on priming effect size showed no significant difference, $t_1(1, 59) < 1$, *ns;* $t_2(1, 11) < 1$, *ns*).[3] As was found for the unambiguous items (86 ms priming in released vs. 81 ms in unreleased condition): the presence of a release burst does not seem to be critical.

Analyses of correct lexical decision rates for all targets with /t/ (as mirror images of the error rates shown in upper part of Table 5) were analysed as well. The factor Ambiguity did not have a significant effect, $F_1(1, 59) < 1$, *ns;* $F_2(1, 34) < 1$, *ns.* Acoustic Variant did have a significant effect on the correct rates, $F_1(3, 57) = 4.04$, $p = .011$; $F_2(3, 32) = 4.25$, $p = .012$; more errors were made in the control conditions than in the test conditions. The interaction between Ambiguity and Acoustic Variant was not significant, $F_1(3, 57) = 2.02$, *ns*; $F_2(3, 32) = 2.50$, $p = .078$.

Response time data for the targets without /t/ (lower part of Table 7) were also subjected to analyses of variance (subjects nested under List in the F_1 analysis). Note that there were only two acoustic variants in this subpart of the design. The effect of Acoustic Variant was highly significant by subjects, but marginally by items, $F_1(1, 54) = 18.05$, $p < .001$; $F_2(1, 11) = 3.70$, $p = .081$. This weak effect in the F_2 analysis may be due mainly to one item yielding a high error rate (the *rust*/*Rus* pair; 'rest/Russian'): *Rus* elicited only 40% correct YES-responses in the 'no /t/' condition.

In sum, when lexical activation is tapped following the licensing phonological context for /t/-deletion, reduced variants and full variants showed equally large priming effects for the intended word ending in /t/. Importantly, however, activation of a target with /t/, given the 'no /t/' acoustic variant, is affected by the presence of a lexical competitor.

This raises the question whether these results provide any evidence for gradient activation. 'Graded activation' was expected, in the sense that the more evidence is found for a certain target word, the stronger the activation of that word. The results of Experiments 1 and 2 have shown that word detection times and repetition priming effects for targets ending in /t/ were

[3] Results in the word detection data were analysed after the exclusion of three items in which the duration of /s/ was unnaturally long. Exclusion of these same three items in this cross-modal study did not affect the pattern of results (significant priming effects in all three test conditions).

equal in released and unreleased conditions. The next section contains a more detailed item analysis to investigate gradiency in the results.

GRADIENT ACTIVATION: COMBINED RESULTS

This section addresses the question whether the results of Experiments 1–3 provide any evidence on gradient activation. Gradient activation implies that variants with more cues for /t/ were more effective in activating the intended word form ending in /t/ than variants with fewer cues for /t/. This was investigated by analysing the unreleased condition items in more detail. So far, results for the released and unreleased conditions in Experiments 2 and 3 have been highly similar, with only weak trends towards higher lexical activation upon presentation of more redundant forms. This suggests that the presence of a release burst itself may not be critical. Nevertheless, there may be variation in the unreleased tokens. Assuming that deletion is actually gradient (cf. Mitterer & Ernestus, 2006), and assuming that there is variation in the unreleased tokens, we might get more insight into 'gradient activation' by comparing the response time results in Experiments 2 and 3 with both the discrimination results from Experiment 1 and with an acoustic measure in the unreleased condition.

The unreleased condition covers at least a part of the continuum between the fully released condition and the 'no /t/' condition. First, the priming results will be compared with the discrimination results. The reasoning was as follows. The easier it is for listeners to hear the difference between the unreleased and 'no /t/' variants (the higher the discriminability), the more cues for /t/ the unreleased form probably contains. Thus, we can compare priming effects for targets with word-final /t/ in items with high discriminability, medium discriminability, and low discriminability (but note that low means *relatively* low: discrimination was high across the board). This was determined for ambiguous items and unambiguous items (again excluding the three unambiguous with relatively long /s/ durations). Discriminability was determined by categorising items in three categories on the basis of percentages correct in the 4-interval oddity task (range is from 67% to 96% correct for both the unambiguous and ambiguous items). The results are given in Table 8.

The results in Table 8 clearly show that the higher the discriminability, the larger the priming effect for words ending in /t/. For both item groups, the correlation strength between priming size (in ms) and discriminability (proportion) was established. The Pearson correlation coefficient was $r = .659$ for the unambiguous items (two-tailed $p < .001$), but correlation was only weak for the ambiguous items ($r = .217$, $p > .1$). When the data of the two item groups are combined, the correlation coefficient was $r = .448$ which

TABLE 8
Priming effect for target with final /t/ broken down by degree of discriminability of the unreleased condition relative to the 'no /t/' condition

	Unambiguous items	*Ambiguous items*
Discriminability	*Mean priming effect*	*Mean priming effect*
High	122 ms ($N = 6/21$)	72 ms ($N = 5/12$)
Medium	81 ms ($N = 10/21$)	52 ms ($N = 4/12$)
'Low'	42 ms ($N = 5/21$)	46 ms ($N = 3/12$)

is significant at the $p < .01$ level. Secondly, the correlation can be established for the discriminability (between the unreleased and 'no /t/' variants in Experiment 1) and the word detection time in the unreleased condition in Experiment 2 for targets with /t/. For both item groups, the correlation strength between detection time (in ms) and discriminability (proportion) was established. The Pearson correlation coefficient was $r = -.076$ for the unambiguous items (*ns*), but there was a significant correlation for the ambiguous items ($r = -.700$, $p = .011$). When the data of the two item groups are combined, the correlation coefficient was $r = -.300$ which is not significant ($p = .090$). These results confirm that it is difficult to find differences between acoustic conditions with the word spotting task: evidence for gradiency in the results of Experiment 2 was found only for the ambiguous items.

Thus, these results provide some indirect support for gradient activation. However, stronger evidence for gradiency would be a direct relation between an acoustic measure and the size of the priming effect in the unreleased condition. Mitterer and Ernestus' study (2006) shows clear examples of cues for /t/ in the absence of a release burst in /st#b/ clusters: fricative /s/ may fade out, followed by a period of low-amplitude frication noise, or there may be rather long (near-) silent interval (too long to be attributed to the following /b/). For each of the 24 unambiguous items, the duration of the silent interval (including the voice bar for /b/) was measured to investigate the correlation between priming effect in the unreleased condition and silent interval duration. This duration varied between 102 and 142 ms (mean of 122 ms). Again, there was a significant relation between this bottom-up cue and lexical activation of a target with /t/ ($r = .526$, $p = .008$ two-tailed). For the ambiguous words, correlation between duration of silent interval (mean duration 113 ms) and activation of the intended word ending in /t/ was weak and insignificant ($r = .185$, $p > .1$). Collapsed over the two item groups, the correlation coefficient was $r = .385$, $p = .021$). There was no significant

correlation between the duration of the silent interval and the word detection results in Experiment 2 ($p > .1$ for both item sets).

Obviously, the speech signal may contain more cues for /t/ than the duration of this silent interval (such as formant movements during /s/ or a certain period of low-amplitude frication). Nevertheless, the combination of the results of Experiment 1 and Experiment 3, and the preliminary acoustic analysis reported here, provide evidence that items in the unreleased condition vary with respect to cues for /t/. The activation level for the intended word is proportional to the amount of bottom-up support for /t/.

Discussion

The results of the cross-modal priming experiment provided clear evidence with respect to the second hypothesis: activation of a target with /t/, given a phonological context that licenses /t/ deletion, is affected by the presence of a lexical competitor matching the surface form. For the ambiguous item set, given the 'no /t/' variant, a priming effect was found for the surface-matching target, but not for the 'reduced' target, even though the target could be activated on the basis of partial overlap. The word detection results (detection rates being particularly compromised for ambiguous items given a 'no /t/' variant) and the results in the present cross-modal study are in line with Gaskell and Marslen-Wilson (2001): context-licensing of a segmental change (be it assimilation or deletion) is strongly decreased if a surface-matching competitor exists.

GENERAL DISCUSSION

Inconsistent results have been reported in the literature with respect to the question whether changed word forms are equally effective in activating the intended word as unchanged word forms. This inconsistency could be due to the fact that different connected-speech processes have been examined. The larger the perceptual difference between the changed and unchanged form, *given the specific phonetic context in which the change occurs,* the larger the effect on lexical activation may be. This study was set up to investigate sound-to-lexicon mapping in the face of (partial) deletion of word-final /t/. The hypothesis was that deletion of /t/ in /st#b/ context is salient, and hence that form variation affects processing time, mainly if it concerns full deletion.

A search through the Spoken Dutch Corpus (Oostdijk, 2000a, 2000b) showed that deletion of noun-final /t/, if defined as the absence of a /t/ release, in the context of /st#b/ was widespread, across speaking styles. Frequency of occurrence of the unreleased variant in this phonological context is thus higher than of the canonical released variant.

A discrimination study showed that listeners could distinguish the three different acoustic versions of each target word (fully released /t/, unreleased /t/, and no /t/) in this specific context. So, despite spectral similarity between /s/ and the release burst of /t/, listeners are nevertheless well able to perceive differences between the different acoustic forms. Note that Mitterer and Blomert (2003) found no Mismatch Negativity component in an oddball experiment for the deviant form *gardem bench,* compared to the standard form *garden bench,* whereas an MMN was observed for the deviant form *gardem chair* (compared with standard *garden chair*). This suggests that deletion of/t/ in this specific /st#b/ context in the present study is more salient than was nasal place assimilation in a viable context in Mitterer and Blomert (2003). Hence, the effect on lexical access was expected to be larger for /t/ deletion than for place assimilation of nasals and plosives.

Lexical activation was investigated in two on-line studies. Despite the fact that listeners could perceive differences between the variants, the results of Experiments 2 and 3 showed that the released and unreleased variants yielded comparable amounts of activation. Thus, importantly, even salient differences do not always affect ease of processing. This is consistent with the findings of Deelman and Connine (2001) and of Sumner and Samuel (2005): regular variants seem to be interchangeable with canonical forms. Nevertheless, closer inspection of the unreleased condition showed that items varied in terms of (the strength of) cues for /t/ and that lexical activation was influenced by the presence and strength of these cues. The amount of evidence for a particular word translates into activation levels in a gradient fashion: the correlation results for the unreleased condition showed this most clearly for the unambiguous items. Therefore, the results are consistent with the hypothesis that form variation affects lexical access: reduced forms are less effective in activating intended targets than more redundant forms. However, the release burst itself hardly seems to contribute to the target's activation: given this phonological context in which most /t/s are unreleased (cf. the Corpus study), the presence of a release burst is not an important cue.

With respect to the full deletion condition, obviously an even more salient change, the results of the cross-modal priming experiment differed from those of the word detection experiment. In the word detection study, the 'no /t/' variant slowed down processing of targets with /t/. In the cross-modal repetition priming experiment, all three variants were equally effective in activating target words with /t/ for the unambiguous items. There may be two explanations for this.

One possible explanation would be low statistical power in the cross-modal study: the priming effect in the released condition (86 ms) was numerically larger than that in the 'no /t/' condition (60 ms), but this difference in effect size failed to reach significance. Because effect size is a derived measure, it may be more difficult to find statistically robust

differences in this measure than in a direct comparison of response times in two conditions. Coenen et al. (2001) encountered similar statistical problems. For regressive coronal assimilation in German (*wort̲ mal* becoming *worp̲ mal*; the changed speech sounds being plosives and nasals), they found numerically less priming for changed (61 ms) than for unchanged words (91 ms). This difference also failed to reach significance.

A second explanation is the difference in activation between the different acoustic forms may have become smaller at the point of visual target presentation. In the word detection study, subjects would press the button as soon as they had received sufficient information. In the 'no /t/' condition, the longer response times for targets with /t/ suggest that listeners waited for the following context to license the variant as an exemplar of a target with /t/. Even though the intended word is activated on the basis of partial overlap up to and including /s/, activation may only surpass a certain threshold once the 'deletion context' is processed. By contrast, in the cross-modal study, the presentation of the visual target was time-locked, in all acoustic variants, to the offset of the following /b/. Measuring from the onset of schwa (or offset of /b/) in the cross-modal study has already allowed the listener to gather more information about /b/ in the 'no /t/' condition, and thus about a possible deleted segment.

Nevertheless, even though activation levels for the intended word in the three acoustic variants may have become more equal at the offset of /b/ than at the onset of /s/, the combined results confirm the hypothesis that form variation affects lexical access, be it that activation differences are relatively small.

The slightly different results in Experiments 2 and 3 with respect to the second hypothesis may also be related to the time course of lexical access. The second hypothesis was that activation of a target with /t/, in the absence of cues for /t/ save a phonological /t/ deletion context, would be affected by the presence of a surface-matching lexical candidate. Although support for this hypothesis was found in the cross-modal priming results of Experiment 3 and in the detection rate analyses in Experiment 2; it is important to note that there is initial bottom-up support for both lexical candidates (cf. e.g., Frauenfelder & Peeters, 1990; McQueen, Cutler, Briscoe & Norris, 1995 on lexical embedding). The high word detection rate (85%) in the compensation condition (listeners detecting *kast* in 'no /t/' conditions) and the high rate of embedded false alarms (higher than 30% when listeners detected pre-assigned *kas* in the released and unreleased /t/ conditions) were explained as an artifact of the word detection task. Pre-assigning a target before presentation of the speech signal in the word detection study may have indeed boosted recognition of this unintended word candidate. Nevertheless, one must assume that both candidates are active at some point, and the question then is *when* the 'reduced candidate' (or the embedded candidate) is

inhibited. The priming results of Experiment 3 have shown that after /b/, the 'reduced' candidate is no longer activated. Word detection *rates* showed a clear interaction: detection rates were particularly compromised in the ambiguous 'no /t/' condition (be it that detection rate was still 85%). The word detection *times* for the ambiguous items in Experiment 2 showed no differences between the acoustic variants and were therefore less clear: responses may have been initiated at an earlier point at which 'reduced candidates' were still activated.

So, the results of the two on-line studies are complementary. Even though Experiment 2 provides evidence both for initial multiple activation and for a bias for the surface-matching candidate, the results obtained at a later point (Experiment 3) confirmed that 'reduced' candidates are disfavoured in semantically neutral contexts. These latter results are therefore consistent with those of Gaskell and Marslen-Wilson (2001) who also tapped activation following the context licensing the segmental change. Gaskell and Marslen-Wilson (2001) found no priming effect for RUN after the presentation of *rum picks* when this sound sequence was embedded in a semantically neutral sentence. However, given the smaller-sized priming effects for the ambiguous items than for the unambiguous items (cf. Table 5), it is clear that lexical competition is not entirely resolved at this point. Therefore, rather than saying that compensation is *blocked* if there is a surface matching candidate, it seems more appropriate to conclude that 'unintended words in ambiguous phrases seem to be an intermediate case: They appear to be activated, but not as strongly as words the speaker intended' (Spinelli et al., 2003, p. 248).

The results therefore fit in with graded activation accounts for spoken word recognition, rather than with a model that makes phonemic decisions at a prelexical stage. Subtle subphonemic cues can be used immediately to increase the activation levels of intended word candidates (Allopenna, Magnuson, & Tanenhaus, 1998; Andruski et al., 1994; Dahan, Magnuson, Tanenhaus, & Hogan, 2001; McMurray, Tanenhaus, & Aslin, 2002; McQueen, Dahan, & Cutler, 2003; Salverda et al., 2003; Spinelli et al., 2003). Activating word candidates on the basis of reduced input, given a viable phonetic context, does slow down the word recognition process because the following context has to be taken into account, even when this segment is not strictly required to distinguish the word from competitors. This contrasts with the results of Deelman and Connine (2001) who investigated lexical activation elicited by illegal (reduced) variants: priming effects in forms in which word-final /s/ was missing (e.g., *bogu_*) were compared with priming effects elicited by intact forms (e.g., intact *bogus*). Deelman and Connine found that word forms without word-final /s/ yielded equally large priming effects as intact words. The discrepancy between the present results and those of Deelman and Connine (2001) might be related to the amount of onset overlap: their stimulus words were all bisyllabic, whereas

words in the present study were monosyllabic. In other words, the relative contribution of the last speech sound to the recognition of the word may be more crucial in the present study.

Thus, form variation may affect lexical activation. However, whether listeners take phonological context into account to enhance activation of potentially reduced word candidates, and whether word-final segments can still contribute to the amount of bottom-up support for a lexical item, is a matter of lexical neighbourhood.

Manuscript received January 2005
Revised manuscript received September 2005
First published online 31 March, 2006

REFERENCES

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.

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

Bard, E. G., Sotillo, C., Kelly, M. L., & Aylett, M. P. (2001). Taking the hit: Leaving some lexical competition to be resolved post-lexically. *Language and Cognitive Processes*, *16*, 731–737.

Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 341–376). Cambridge: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155–180.

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

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Cambridge, MA: MIT Press.

Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes*, *16*, 535–564.

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

Cutler, A., & Butterfield, S. (1990). Word boundary cues in clear speech. *Speech Communication*, *9*, 485–495.

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

Deelman, T., & Connine, C. M. (2001). Missing information in spoken word recognition: Nonreleased stop consonants. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 656–663.

Ellis, L., & Hardcastle, W. J. (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics*, *30*, 373–396.

Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch: A corpus-based study of the phonology-phonetics interface*. Doctoral dissertation Vrije Universiteit Amsterdam. Utrecht: LOT series.

Ernestus, M., Baayen, R. H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language*, *81*, 162–173.

Frauenfelder, U. H., & Peeters, G. (1990). Lexical segmentation in TRACE: An exercise in simulation. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistics and computational perspectives* (pp. 50–86). Cambridge, MA: MIT Press.

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.

Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 380–396.

Gaskell, M. G., & Marslen-Wilson, W. D. (2001). Lexical ambiguity resolution and spoken word recognition: Bridging the gap. *Journal of Memory and Language*, *44*, 325–349.

Gerrits, E. (2001). *The categorisation of speech sounds by adults and children*. Doctoral dissertation Utrecht University. Utrecht, the Netherlands: LOT.

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

Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, *45*, 133–159.

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

Gow, D. W., & Im, A. M. (2004). A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language*, *51*, 279–296.

Guy, G. R. (1980). Variation in the group and in the individual: The case of final stop deletion. In W. Labov (Ed.), *Locating language in time and space* (pp. 1–36). New York: Academic Press.

Hura, S. L., Lindblom, B., & Diehl, R. L. (1992). On the role of perception in shaping phonological assimilation rules. *Language and Speech*, *35*, 59–72.

Janse, E. (2004). Word perception in fast speech: Artificially time-compressed vs. naturally produced fast speech. *Speech Communication*, *42*, 155–173.

Kaplan, H. L., Macmillan, N. A., & Creelman, C. D. (1978). Methods and designs. *Behavior Research Methods and Instrumentation*, *10*, 796–813.

Kemps, R., Ernestus, M., Schreuder, R., & Baayen, R. H. (2004). Processing reduced forms: The suffix restoration effect. *Brain and Language*, *90*, 117–127.

Kilborn, K., & Moss, H. (1996). Word monitoring. *Language and Cognitive Processes*, *11*(6), 689–694.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Kohler, K. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 69–92). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Koster, C. J. (1987). *Word recognition in foreign and native language*. Dordrecht, the Netherlands: Foris.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*, *51*, 2018–2024.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

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

McQueen, J. M., Cutler, A., Briscoe, T., & Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes*, *10*, 309–331.

McQueen, J. M., Dahan, D., & Cutler, A. (2003). Continuity and gradedness in speech processing. In N.O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 39–78). Berlin: Mouton de.

Mitterer, H. (2003). *Understanding ''gardem bench'': Studies on the perception of assimilated speech*. Doctoral dissertation University of Maastricht, the Netherlands.

Mitterer, H., & Ernestus, M. (2006). Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics*, *34*, 73–103.

Mitterer, H., Csépe, M., & Blomert, L. (2003). Compensation for phonological assimilation in perception: Evidence from Hungarian liquid assimilation. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, 2321–2324.

Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception and Psychophysics*, *65*(6), 956–969.

Mitterer, H., & Ernestus, M. (2006). Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics*, *34*, 73–103.

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, *125*, 826–859.

Neu, H. (1980). Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. In W. Labov (Ed.), *Locating language in time and space* (pp. 37–54). New York: Academic Press.

Nolan, F. (1992). The descriptive role of segments: Evidence from assimilation. In G. J. Doherty & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, segment and prosody* (pp. 261–280). Cambridge: Cambridge University Press.

Ohala, J. J. (1990). The phonetics and phonology of aspects of assimilation. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 258–275). Cambridge: Cambridge University Press.

Oostdijk, N. (2000a). The Spoken Dutch Corpus. Overview and first evaluation. In M. Gravilidou, G. Carayannis, S. Markantonatou, S. Piperidis & G. Stainhaouer (Eds.), *LREC-2000 (Second International Conference on Language Resources and Evaluation) Proceedings*. Vol. II: 887–894.

Oostdijk, N. (2000b). Het Corpus Gesproken Nederlands. *Nederlandse Taalkunde*, *5*(3), 280–284.

Otake, T., Yoneyama, K., Cutler, A., & van der Lugt, A. (1996). The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America*, *100*, 3831–3842.

Pickett, J. M., & Pollack, I. (1963). Intelligibility of excerpts from fluent speech: Effects of rate of utterance and duration of excerpt. *Language and Speech*, *6*, 151–164.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 785–794.

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.

Schouten, B., Gerrits, E., & van Hessen, A. (2003). The end of categorical perception as we know it. *Speech Communication*, *41*, 71–80.

van Son, R. J. J. H., Binnenpoorte, D., van den Heuvel, H., & Pols, L. C. W. (2001). The IFA Corpus: A phonemically segmented Dutch ''open source'' speech database. *Proceedings Eurospeech 2001, Aalborg, Denmark*, Vol. 3, 2051–2054.

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

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language*, *52*, 322–338.

Turk, A. E., & Shattuck-Hufnagel, S. (2000). Word-boundary-related duration patterns in English. *Journal of Phonetics*, *28*, 397–440.

Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General*, *123*, 34–80.

Zwaardemaker, H., & Eijkman, L. P. H. (1928). *Leerboek der phonetiek*. Haarlem: De Erven F. Bohn.

Zwitserlood, P., & Schriefers, H. (1995). Effects of sensory information and processing time in spoken-word recognition. *Language and Cognitive Processes*, *10*, 121–136.

APPENDIX

Ambiguous target items

1. Hij kan beter de *kas*/*kast* bewaken.
2. Dorine wou hem geen *haas*/*haast* bezorgen.
3. Jan wil graag de *bas*/*bast* bewerken.
4. Ik zou niet meer weten hoe ik het woord '*gas*/*gast*' beschreven heb.
5. Kon hij niet eens het woord '*kers*/*Kerst*' beschrijven?
6. Hij wilde de man geen *mes*/*mest* beloven.
7. Je kunt over die *hoes*/*hoest* beweren wat je wilt.
8. Hoe wou je dan die *roes*/*roest* beperken?
9. En hoe ga je de *Rus*/*rust* bewaken?
10. Daarna heeft hij de *mis*/*mist* beschreven.
11. Ik zal me niet meer met de *reis*/*rijst* bemoeien.
12. Ze wilden het niet tot de *kus*/*kust* beperken.

Unambiguous target items

1. Nu moest hij nog de *rest* bevestigen.
2. Of hij werkelijk zo slim is, dat moet de *test* bewijzen.
3. Maar je kunt toch geen *geest* bestraffen?
4. Hem *last* bezorgen is wel het laatste wat ik wil.
5. Je kon de *mast* bewegen, zo krakkemikkig was het schip.
6. Je kunt maar beter de *post* bewaren.
7. Je kunt de *lijst* bestellen bij het landelijk bureau.
8. Martin kon in z'n eentje de *kist* bewegen.
9. Als jij nu eens een *list* bedenken kon . . .
10. Maria kon geen *gist* betalen.
11. In de grote steden kon men misschien de *pest* bedwingen, (maar niet de andere besmettelijke ziektes)
12. Kunnen we nog even dat *nest* bezoeken waar die zwaan op haar eieren zit?
13. De peuter wilde liever het *beest* bezoeken, (want oma zag hij al elke week.)
14. Hij kon het *feest* bekostigen, (maar daar was alles mee gezegd.)
15. Ze kon haar *dorst* bestrijden met de grote glazen frisdrank die al klaar stonden.
16. Ik wist niet dat hij een druk op de *borst* bedoelde.
17. Je gaat iemand toch niet van z'n *vest* beroven?
18. Ik zou die *puist* bedekken, als ik jou was.
19. Hij wilde hem met z'n *vuist* bewerken, (maar daar stak de politie een stokje voor.)
20. Peter heeft de *korst* bedekt met een pleister.
21. Misschien kunnen we daar wel die *kwast* bestellen.
22. Thomas wilde per se de *worst* bezorgen bij het bejaardentehuis.
23. Hij zal toch niet de *herfst* bedoelen?
24. Vreemd genoeg kon hij wel de *vorst* bereiken op zijn vakantieadres.