# Recognizing reduced forms: Different processing mechanisms for similar reductions

Holger Mitterer *

*Max Planck Institute for Psycholinguistics, Postbus 310, 6500 AH Nijmegen, The Netherlands*

## ARTICLE INFO

## ABSTRACT

Recognizing phonetically reduced forms is a huge challenge for spoken-word recognition. Phonetic reductions not only occur often, but also come in a variety of forms. The paper investigates how two similar forms of reductions – /t/-reduction and nasal place assimilation in Dutch – can eventually be recognized, focusing on the role of following phonological context. Previous research indicated that listeners take the following phonological context into account when compensating for /t/-reduction and nasal place assimilation. The current paper shows that these context effects arise in early perceptual processes for the perception of assimilated forms, but at a later stage of processing for the perception of /t/-reduced forms. This shows first that the recognition of apparently similarly reduced words may rely on different processing mechanisms and, second, that searching for dissociations over tasks is a promising research strategy to investigate how reduced forms are recognized.

## 1. Introduction

Phonological reductions are not only ubiquitous, they are also quite diverse. Dilley and Pitt (2007) showed that a word-final /t/ can be reduced in at least three different ways. It can surface as a /p/ (e.g., *got bored* /gɑtboʊɹd/→[gɒpboʊɹd], due to place assimilation by the following /b/), as a glottal stop ([gɑʔboʊɹd]), or be absent altogether ([gɑboʊɹd]). Moreover, none of these alternatives is completely marginal, so that all forms will occur regularly in spontaneous speech.

From the point of view of spoken-word recognition, the question arises how words can still be recognized when they have segments missing or altered. The goal of this paper is to argue that there is something wrong with asking the question in this way. The question implies that generalization from one phonetic reduction process to another is straightforward. Such a generalization may seem reasonable if the production patterns are similar for two forms of phonetic reduction. This is the case for assimilation of word-final /n/ and the reduction and deletion of word-final /t/ in Dutch. Both can be triggered by a following /b/: A word-final /n/ may then become an [m] (cf. Booij, 1995) and a word-final /t/ is more likely to be reduced in this than in any other following context (cf. Mitterer & Ernestus, 2006). Reduction of /t/ here means that the release burst (which is typically present in Dutch stops, cf. Cho & McQueen, 2008) is missing. An additional parallel of these two forms of phonetic reduction is that both can be explained by an overlap of the /b/'s labial closing gesture with the alveolar closing gestures for /n/ and /t/. In case of nasal place assimilation, this leads to a nasal murmur with a labial closure; and in the case of /t/ reduction, the alveolar release becomes inaudible because of the overlapping labial closure. While differences remain between /t/-reduction and nasal place assimilation – for instance, /t/-reduction is also strongly influenced by preceding context and nasal place assimilation is not – the two processes are similar in that they are both facilitated by the presence of a following /b/, and this facilitation can be explained by gestural overlap. The purpose of this paper is to show, however, that, in perception, the effect of following context is triggered by different processing mechanisms.

In perception, following context is nevertheless taken into account in both cases. Gaskell and Marslen-Wilson (1996) used cross-modal priming and found that an assimilated form (e.g., [lim], from *lean*) primed recognition of the visual word *lean* if it was presented in a labial context that triggers assimilation but not if it was presented in an unviable context for assimilation. For /t/-reduction, Mitterer and McQueen (2009) found that listeners were more likely to interpret a form such as [kys] as an instance of the Dutch word *kust* /kyst/ (Engl., 'coast') and not the word *kus* /kys/ (Engl., 'kiss') if the form is presented in a labial context that makes the /t/-reduction likely.

It seems prima vista unlikely that the spoken-word recognition system would invent the wheel twice. Positing different processing mechanisms for these two effects triggered by the following context seems like a prime example of "needlessly multiplying entities". However, if data forced us to conclude that there are different mechanisms at work, these data would provide a need to multiply entities. The purpose of this paper is to present such data.

* Tel.: +31 24 3521375; fax: +31 24 3521213.
  *E-mail address:* holger.mitterer@mpi.nl

Additional research on the perception of assimilated forms showed that the effect of following context is very robust. In addition to cross-modal priming and phoneme monitoring (Coenen, Zwitserlood, & Bölte, 2001; Gaskell & Marslen-Wilson, 1996, 1998, 2001; Gow, 2002, 2003), effects of following context have been observed in an eye-tracking task (Gow & McMurray, 2007), in identification tasks, and passive electrophysiological recordings (Mitterer & Blomert, 2003). The latter effects suggested that the compensation effect may be related to compensation for coarticulation. Although there is a lively debate about the causes of compensation for coarticulation (Fowler, 1996, 2006; Fowler, Brown, & Mann, 2000; Fowler & Smith, 1986; Lotto & Kluender, 1998), there is no disagreement that it arises as an automatic consequence of hearing speech sounds, and models that assume several stages in the perceptual process, such as the general-auditory framework (Diehl, Lotto, & Holt, 2004), assume a locus at an early stage.

One aspect of compensation for coarticulation is that it does not depend on specific language experience (Lotto, Kluender, & Holt, 1997; Mann, 1986). Mitterer and colleagues hence tested the role of experience in compensation for assimilation (Mitterer, Csépe, & Blomert, 2006; Mitterer, Csépe, Honbolygo, & Blomert, 2006). They used a Hungarian liquid assimilation rule (/l/ before an /r/ can surface as an [r], /bɑlrol/→[barrol]) to test compensation by native Hungarian and non-native (Dutch) listeners. This form of assimilation does not occur in Dutch (not even in casual speech, cf. Ernestus, 2000), so that this assimilation was unfamiliar for Dutch listeners. The results showed that, nevertheless, both Hungarian and Dutch listeners compensated for this assimilation with context effects in an identification task, a discrimination task, and in passive electrophysiological recordings. It has to be noted that the Dutch listeners did not show exactly the same data pattern as the Hungarian listeners. This indicates that there are some language-specific, learning-dependent contributions to compensation for assimilation, in line with results from other labs (Darcy, Peperkamp, & Dupoux, 2007; Snoeren, Gaskell, & DiBetta, 2009). Darcy et al., for instance, showed that the degree to which listeners compensate for French voicing assimilation is correlated with their command of French. English monolinguals showed only a small effect, French native speakers showed a large effect, and English learners fell in between these two groups. Nevertheless, the overall data pattern of Dutch and Hungarian listeners over all tasks in the papers by Mitterer and co-workers (Mitterer, Csépe, & Blomert, 2006; Mitterer, Csépe, Honbolygo, et al., 2006) showed more commonalities than differences. Importantly, both groups of listeners showed a context effect in a discrimination task: The difference between the canonical and the assimilated form was less salient in the context that allowed assimilation. This showed that learning-independent accounts assuming an early perceptual locus of the effect – as proposed for compensation for coarticulation – are tenable for compensation for assimilation as well.

The rationale in these series of experiments was the following: If compensation for assimilation is triggered by early and mandatory perceptual processes, it should be difficult to find a dissociation so that the context effect arises in one task but not in another. Despite trying many different tasks and measures – identification with and without feedback, discrimination, electrophysiological recordings, using native and non-native listeners, and using speech and non-speech contexts – the context effect in compensation for assimilation turned out to be stable. Based on these results, Mitterer, Csépe, and Blomert (2006) argued that early perceptual processing triggers compensation for assimilation, similar to compensation for coarticulation. This dovetails well with proposals from linguistics that assimilation rules are shaped by the sensitivities (or better: lack of sensitivity) of the listener (Hura, Lindblom, & Diehl, 1992; Steriade, 2001). Assimilations tend to occur in environments in which their consequences are difficult to hear, so that the assimilated form sounds similar to the canonical form. This clearly seems to be the case for nasal place assimilation; Ohala (1975, 1990) argued that nasals are highly distinct from other consonants but very similar to each other. Hence, a change in place of articulation is difficult to notice in a nasal.

Turning to /t/-reduction, the question arises as to whether the effect of following context is also found in many different tasks for this reduction. It seems not, given that Mitterer and McQueen (2009) found a context effect using an eye-tracking measure, but Mitterer and Ernestus (2006) failed to find an effect in an identification task. That is, listeners were not more likely to "reconstruct" a phonetically reduced /t/ if the following context was a /b/ than if it was a /k/.

Together with the positive results in Mitterer and McQueen (2009), there is hence a dissociation of the context effect over tasks. The context effect did not arise in the identification task, but did arise in the eye-tracking experiment. This vindicates the research strategy proposed above, in which much of the argument rests on not finding a dissociation over tasks. The fact that no such dissociation was found for compensation for assimilation is hence not due to such dissociations being impossible.

On a broader scale, this also suggests that compensation for /t/-reduction and compensation for assimilation are not triggered by the same processing mechanisms. Otherwise, we would have observed the same pattern of context effects over different tasks. Accordingly, this seems to support the point made at the beginning of this paper: the recognition of different reduced word forms depends on different processing mechanisms. Whatever processing mechanism triggers compensation for assimilation at an early perceptual level of processing seems not to be involved in compensation for /t/-reduction.

There remain, however, two reservations. First, the chain of arguments presented here is to some extent internally inconsistent. Even though it was claimed that one should not generalize from one type of phonetic reduction process to another, I did just that by generalizing from a Hungarian liquid assimilation rule to nasal place assimilation. The results reported above only showed that the consequences of Hungarian place assimilation are inconsequential for word recognition. Based on the argumentation developed here, this does not mean that this is also true for nasal place assimilation in Dutch. Different assimilations may after all have rather different acoustic consequences, leading to different forms of processing (see Gow & Im, 2004). So it remains unclear whether the compensation for nasal place assimilation relies on early perceptual processes as well.

A second reservation is that the chain of arguments depends crucially on the one null effect in the identification experiment presented by Mitterer and Ernestus (2006). In this experiment, the following context did not influence listeners' decisions on whether or not a word contained a word-final /t/. There were significant effects of preceding context and phonetic detail in this experiment, which makes the null effect of following context more meaningful. Nevertheless, the empirical basis remains weak for the case that the effect of following context on compensation for /t/-reduction does not arise in early perceptual processes.

Both shortcomings are addressed here in a single experiment. The experiment is set up to show that the perceptual salience of nasal place assimilation depends on the following context, while the perceptual salience of /t/-reduction does not. As in the research on assimilation and /t/-reduction described above, this was tested with the four-interval (4I) oddity task. In this task, participants hear four stimuli, of which either the second or the third is different from the other three. After hearing the series of four stimuli, the participant has to indicate whether the second or the third one is the "odd one out". The choice for this task was based on work by

Gerrits and colleagues (Gerrits & Schouten, 2004; Schouten, Gerrits, & van Hessen, 2003). They tested a range of discrimination tasks and found that "categorical perception effects" in discrimination performance (i.e., a performance peak at the category boundary) were entirely dependent on the task used. In the 4I-oddity task, there was no influence of phonological categorization on discrimination performance. That is, between-category differences were not easier to discriminate than within-category differences. This indicates that the 4I-oddity task reveals how well two stimuli are discriminable on an auditory level, without an influence of phonological categorization.

With this task, the predictions are the following: The salience of the difference between a canonically produced /n/-final word (tuin /tœyn/, Engl. 'garden') and an assimilated version ([tœym]) is reduced in a viable context that allows the assimilation in comparison to an unviable context that does not allow the assimilation. In contrast, the salience of the difference between a word with a word-final /t/ produced canonically (kust /kyst/, Engl. 'coast') and a /t/-reduced variant ([kys]) with no burst is the same in a viable context that facilitates phonetic reduction (i.e., /b/) and an unviable context that does not facilitate phonetic reduction (i.e., /n/).

## 2. Method

### 2.1. Participants

Twenty members of the subject pool of the Max Planck Institute for Psycholinguistics participated in the experiment. All participants were native speakers of Dutch, had no history of hearing problems, and were paid for their participation.

### 2.2. Stimuli

The stimuli were adapted versions of the stimuli used by Mitterer and Blomert (2003) and Mitterer and McQueen (2009). In both experiments, there were context effects for compensation for assimilation and /t/-reduction, respectively. This means that an absence of context effects in the current experiment cannot be attributed to the stimuli being inadequate to produce context effects.

The stimuli for the assimilation condition were constructed on the basis of the stimuli used by Mitterer and Blomert (2003). The original stimuli consisted of two-syllable words and were truncated after the second vowel for the current experiment. The participants therefore had to discriminate [tœynbɑ] from [tœymbɑ] and [tœynstu] from [tœymstu]. Because the f0 was lower in the [m] than in the [n] in the original stimuli – and this could be used in a discrimination task – the stimuli were re-synthesized (using PSOLA) with an identical, linearly falling f0 contour, using the minimum (60 Hz) and maximum (90 Hz) in the original recordings.

The stimuli for the /t/-reduction condition were taken from the study of Mitterer and McQueen (2009). They used 32 minimal pairs in two experiments. From these, the pair kust/kus /kys(t)/ (Engl., 'kiss' and 'coast') was selected because it gave rise to an above-average context effect in both experiments in that study. To affirm the context effect for this item, the effect was tested for statistical significance using the data set of Mitterer and McQueen (2009). Although that experiment was not set up to show effects of single items, a linear-mixed effect model based on only the data with this one item – less than 4% of the complete data set – with context (boven versus naast) as fixed effect and participant as a random factor showed a significant effect of context ($t(93)=2.28, p < 0.05$).

To have a context of a similar duration as in the assimilation condition, we also used only the CV part of the following words

boven /bovə/ and naast /nast/ (Engl., 'above' and 'next to'). Consequently, the pairs to be discriminated were [kstbo] versus [kysbo] and [kystnɑ] versus [kysnɑ]. Note that the stimuli without a /t/ still contained a voiceless closure; the transcription without a /t/ is a convenient approximation for presentational purposes.

### 2.3. Procedure and design

The design entails two factors with two levels each. The first factor is Reduction Process (levels: assimilation vs. /t/-reduction) and the second factor is Context (levels: viable vs. unviable for the reduction to occur). The viable context is the /b/-initial syllable for both the assimilation and the /t/-reduction condition, and the unviable context is [na] for the /t/-reduction condition and [stu] for the assimilation condition.

The two conditions "assimilation" and "/t/-reduction" were blocked. Within each block, the context condition was mixed. Each block consisted of 96 trials, with 48 presentations of each context. On a given trial the participants heard four stimuli, three of which were the canonical version and one that was the reduced version. The reduced version was either the second or the third stimulus of the four, and the participants had to indicate whether the second of the third stimulus was different from the rest. The odd stimulus was 24 times in the second and 24 times in third position in each context condition.

Before each block, participants were briefly familiarized with the two stimuli out of context to focus listeners' attention on the relevant part of the stimulus in which a change could occur. Pretesting had shown that participants were otherwise easily frustrated because they did not know in which part of the stimulus a change could occur. The order of blocks was counterbalanced across participants.

## 3. Results

The mean percentages of correct response (p(c)) per condition is displayed in the left panel of Fig. 1. Because the individual p(c) were transformed logistically for the statistical analysis, the mean of these transformed p(c) are displayed in the right panel of Fig. 1. Participants performed at chance level in the viable assimilation condition (51% correct), better than chance in the unviable assimilation condition (64% correct), and well above chance level in the viable and unviable /t/-reduction conditions (83% and 80% correct, respectively). The logistically transformed p(c) were used in a two-factorial repeated measurement ANOVA, which revealed a significant effect of Reduction Process ($F(1, 19)=42.36, p < 0.001$), a marginal effect of Context ($F(1, 19)=4.12, p=0.057$), and, most importantly, an interaction of these two factors ($F(1, 19)=13.05, p < 0.005$). Because of the significant interaction, the effect of Context was tested for both reduction processes separately. This revealed an effect of context for the assimilation condition ($t(19)=4.70, p < 0.001$), but not for the /t/-reduction condition ($t(19)= -1.14, p > 0.1$). Note that this non-significant context effect for the /t/-reduction condition has a different sign than the context effect for the assimilation condition: in the /t/-reduction condition, performance was slightly better in the viable condition than in the unviable condition.

## 4. Discussion

The context effect for the perception of assimilated forms in a discrimination task (Mitterer, Csépe, & Blomert, 2006) was replicated here. But no context effect was found when listeners discriminated canonical and reduced variants of words with
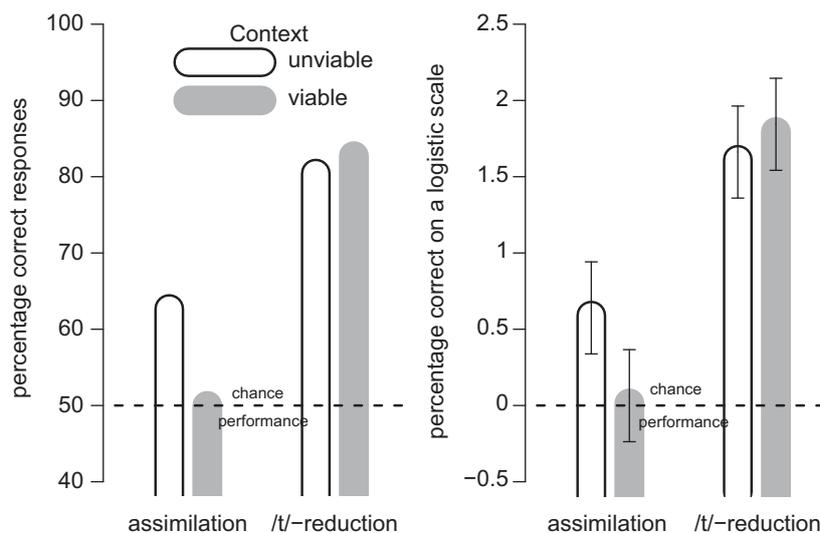
**Fig. 1.** Mean discrimination performance as percentage correct and logistically transformed percentages correct in the four conditions of the experiment. Error bars indicate confidence intervals (CI) around the mean, based on the error terms of the ANOVA with the logistically transformed percentages of correct responses (Masson & Loftus, 2003).

word-final /t/. These results show a dissociation of the context effect within one experiment, with a context effect in the assimilation condition but not in the /t/-reduction condition. Additionally, the results show a dissociation for the case of /t/-reduction over different experiments. The item, which gave rise to a significant context effect in the data of Mitterer and McQueen (2009), failed to give rise to a context effect in the current experiment. This seems to indicate that compensation for assimilation is already triggered at an early perceptual level while compensation for /t/-reduction is not.

Before we can accept this conclusion, there are possible alternative explanations to consider. As pointed out by an anonymous reviewer, the stimuli in the assimilation condition and in the /t/-reduction condition differed in their lexical status. Both stimuli in the stimuli in the /t/-reduction condition are existing words in Dutch, but, in the assimilation condition, only the canonical stimulus (*tuin*), but not the reduced stimulus (*tuim*), is a word in Dutch. This difference in lexical status is, however, an unlikely cause of the observed results. Mitterer, Csépe, and Blomert (2006) manipulated lexical status in the same discrimination task used here and observed no effect of lexical status whatsoever.

A second possible alternative explanation, also suggested by an anonymous reviewer, stems from the fact that the stimuli in the /t/-reduction condition were overall easier to discriminate than the stimuli in the assimilation condition. It should be noted that this is not by itself an explanation of the different effects of context; performance in the assimilation condition could have been poor overall and not be influenced by context at all. However, this difference is still technically a confound; it may just be easier to find context effects near chance level performance. However, the results of Mitterer, Csépe, and Blomert (2006), Mitterer, Csépe, Honbolygo, et al. (2006) make this unlikely as well. In these experiments, the context effect in compensation for assimilation was smaller when discrimination performance was near chance level in the unviable baseline condition. If anything, it is hence more difficult to find context effects near chance performance in this task, so that the overall difference in discriminability between the assimilation condition and the /t/-reduction condition, if anything, worked against the obtained finding.

Even though the baseline difference in discriminability does not explain the difference between the two items on task-related grounds, it may provide a hint as to why there may be differential involvement of early perceptual processes in compensation for different reductions. Apparently, if the perceptual difference between a full and a reduced form becomes more audible, the less likely it is that compensation can already be triggered at an early level of processing. This distinction is to some extent mirrored in functional theories of phonology, which view phonology as a compromise between a speaker who wants to "get away" with as little articulatory effort as possible, while the listener wants a signal that is as clear as possible. In this interplay, some reductions may arise because the listener is not bothered by the consequences, and nasal place assimilation may be such a listener-driven reduction. In other cases, the articulatory effort for the canonical form would be so high, that the speaker reduces regardless of the perceptual consequences. The realization of a /tb/ sequence might be such a case, as it would necessitate achieving two closures and their releases sequentially in a short time window. Note that such patterns can arise in a memetic evolution of language (Blevins, 2004) without the conscious deliberation of either speaker or listener.

A final caveat is the possibility that the dissociation is not so much between /t/-reduction and place assimilation, but rather between the tokens used in this study. Obtaining a reliable measure of perceptual difference requires about 20 repetitions per stimulus, so that it is not feasible to test a large range of stimuli without creating a prohibitively long experimental session. Being restricted to two pairs of stimuli, it remains possible that the dissociation is not so much related to the processes of /t/-reduction and place assimilation, but rather to the specific instances used in this study. Gaskell and Snoeren (2008) found that assimilations vary in strength, and Mitterer and Ernestus (2006) reported similar gradations for /t/-reductions. The stimuli used in this study are cases of a weak form of assimilation and a quite strong form of /t/-reduction. This is also reflected in the baseline differences in discriminability in the current data. It hence remains possible that the differences were to due to "reduction strength" rather than the specific reduction process. A modeling study by Gaskell (2003) showed that assimilation strength may influence how compensation is achieved. A similar result as obtained for /t/-reduction might have been obtained with a more strongly assimilated token of a nasal. It worth noting, however, that this caveat nevertheless underscores the main point of this paper: similar reductions may be processed differently, so that even two tokens of place assimilation might be compensated for differently depending on their assimilation strength.

In line with this assumption, it has been shown that compensation for assimilation is not solely due to early perceptual processes, but is enhanced by higher-level factors such as familiarity with the assimilation (Darcy et al., 2007; Mitterer, Csépe, & Blomert, 2006) and the lexical status of the assimilated words (Gaskell & Marslen-Wilson, 1998; Snoeren et al., 2009). This enhancement may be achieved by the same phonological processing mechanisms which mediate the effect of following context in the perception of /t/-reduction.

Nevertheless, the current data show that similar reductions are not necessarily compensated for in a similar fashion. Even though the current data do not allow us to pinpoint the source of the dissociation – reduction process or reduction strength – they show that early perceptual processes help to compensate for some reductions but not for others. This contrasts with several models which assume some generality in the perception of phonological variants. Proponents of exemplar models of spoken-word recognition (Bybee, 2001; Hawkins, 2003; Johnson, 1997; Pierrehumbert, 2002) argue that all variations are stored in the lexicon. When the listener then encounters a reduced word in the input, this word form will match an exemplar of another reduced variant encountered previously. Word recognition then proceeds via the activation of a variant form in the lexicon. There is evidence for this claim with regard to flapping of alveolar stops (Connine, 2004; Pitt, 2009) and schwa-reduction after strong syllables (Connine, Ranbom, & Patterson, 2008) in American English. This account, in principle, allows the recognition of any reduced form. A second class of theories that can potentially explain a wide variety of context effects in speech perception are gestural theories such as the direct-perception account (Fowler, 1996; Fowler & Smith, 1986; Goldstein & Fowler, 2003) and Motor Theory (for a recent review, see Galantucci, Fowler, & Turvey, 2006). These theories argue that listeners deconvolve the overlapping speech gestures into their sound-producing gestures, which could explain a wide range of compensation for coarticulation findings as well as compensation for assimilation and /t/-reduction, which can also be viewed as caused by gestural overlap. Finally, the statistical learning account proposed by Gaskell (2003) also lends itself to explain a wide variety of context-sensitivities in the perception of speech.

However, the current data indicate that no single mechanism can explain both compensation for assimilation and /t/-reduction sufficiently. A given model may be able to account for any given pattern of effects, that is, why the context influences performance on task A but not on task B. These task effects would have to be consistent, though, and not vary with the reduction process under study. To give an example: from the point of the Motor Theory of speech perception, one may argue that context effects arise only in the speech mode of perception, but not in the non-speech mode. This may explain the dissociation observed for the effect of following context on the perception of /t/-reduction. In the eye-tracking, word-recognition task, participants operate in speech mode; but, in a discrimination task, they operate in a non-speech mode. While this explains the effects observed for /t/-reduction, it becomes difficult to argue that the same mechanism accounts for the effects observed for assimilations: in this case, discrimination should also be unaffected by context while eye-tracking results should show context effects.

The point of this example is not to discredit Motor Theory as a whole. In fact, the point is not to downplay the importance of any of mechanisms just sketched for the perception of reduced forms. The current data only question the generality of any mechanism. Because reductions come in different forms and strengths, perception has to use different strategies to recognize different reductions. Still, each of these mechanisms may play an important role in compensation for some – but clearly not all – types of phonological reduction. The point to be made here is a more general one: no single mechanism can account for the recognition of the large variety of reduced forms. More specifically, even if two reduction processes are due to gestural overlap with a following /b/ – like nasal place assimilation and /t/-reduction – one cannot safely assume that processing mechanisms that mediate a context effect in perception for one type of reduction will also account for the context effect in the perception of the other form of reduction.

A second theoretical consequence of the current result is that looking for dissociations is a valuable tool to understand the processing mechanisms that underlie the perception of reduced forms. The argument for an early locus of compensation for assimilation (Mitterer, Csépe, & Blomert, 2006) was based on the absence of a dissociation over tasks. Such a dissociation has now firmly been established for the case of compensation for /t/-reduction. This shows that our initial quest for a dissociation in the case of compensation for assimilation was not a wild-goose chase. Moreover, there now exists a wide range of methods to investigate the mechanisms that allow the recognition of phonological variants. In addition to the methods investigating perceptual mechanisms presented here, Connine and co-workers (Connine, 2004; Connine et al., 2008), as well as Pitt (2009), present some rather creative methods to investigate the role of lexical storage in the recognition of reduced forms. The task of future research is therefore to test a wide variety of reduction processes with this wide variety of methods. This alone will allow us to understand how auditory and phonological processing as well as lexical storage contribute to the recognition of reduced forms.

## References

Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

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

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

Cho, T., & McQueen, J. M. (2008). Not all sounds in assimilation environment are perceived equally: Evidence from Korean. *Journal of Phonetics*, 36, 239–249.

Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Consequences of assimilation for word recognition and lexical representation. *Language and Cognitive Processes*, 15, 535–564.

Connine, C. M. (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin and Review*, 11, 1084–1089.

Connine, C. M., Ranbom, L. J., & Patterson, D. J. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics*, 70, 403–411.

Darcy, I., Peperkamp, S., & Dupoux, E. (2007). Bilinguals play by the rules: Perceptual compensation for assimilation in late L2-learners. In J. Cole, & J. Hualde (Eds.), *Laboratory phonology 9* (pp. 411–442). Berlin: Mouton de Gruyter.

Diehl, R., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology*, 55, 149–179.

Dilley, L., & Pitt, M. A. (2007). A study of regressive place assimilation in spontaneous speech and its implications for spoken word recognition. *Journal of the Acoustical Society of America*, 122, 2340–2353.

Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch: A corpus-based study of the phonology–phonetics interface [dissertation]*. Utrecht, the Netherlands: LOT.

Fowler, C. A. (1996). Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America*, 99, 1730–1741.

Fowler, C. A. (2006). Compensation for coarticulation reflects gesture perception, not spectral contrast. *Perception & Psychophysics*, 68, 178–183.

Fowler, C. A., Brown, J. M., & Mann, V. A. (2000). Contrast effects do not underlie effects of preceding liquids on stop-consonant identification by humans. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 877–888.

Fowler, C. A., & Smith, M. (1986). Speech perception as "vector analysis": An approach to the problems of segmentation and invariance. In J. Perkell, & D. Klatt (Eds.), *Invariance and variability of speech processes* (pp. 123–136). Hillsdale, NJ: Lawrence Earlbaum Associates.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin and Review*, 13, 13.

Gaskell, M. G. (2003). Modelling regressive and progressive effects of assimilation in speech perception. *Journal of Phonetics*, 31, 447–463.

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.

Gaskell, M. G., & Snoeren, N. D. (2008). The impact of strong assimilation on the perception of connected speech. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1632–1647.

Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, 66, 363–376.

Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In N. O. Schiller, & A. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 159–207). Berlin: Mouton de Gruyter.

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. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics*, 65, 575–590.

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

Gow, D. W., & McMurray, B. (2007). Word recognition and phonology: The case of English coronal place assimilation. In J. Cole, & J. Hualde (Eds.), *Laboratory phonology 9* (pp. 173–200). New York: Mouton de Gruyter.

Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics*, 31, 373–405.

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

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson, & J. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego, CA: Academic Press.

Lotto, A. J., & Kluender, K. R. (1998). General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification. *Perception & Psychophysics*, 60, 602–619.

Lotto, A. J., Kluender, K. R., & Holt, L. L. (1997). Perceptual compensation for coarticulation by Japanese quail (Coturnix coturnix japonica). *Journal of the Acoustical Society of America*, 102, 1134–1140.

Mann, V. A. (1986). Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners perception of English 'l' and 'r'. *Cognition*, 24, 169–196.

Masson, M. E., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, 57, 203–330.

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

Mitterer, H., Csépe, V., & Blomert, L. (2006). The role of perceptual integration in the perception of assimilated word forms. *Quarterly Journal of Experimental Psychology*, 59, 1395–1424.

Mitterer, H., Csépe, V., Honbolygo, F., & Blomert, L. (2006). The recognition of assimilated word forms does not depend on specific language experience. *Cognitive Science*, 30, 451–479.

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

Mitterer, H., & McQueen, J. M. (2009). Processing reduced word-forms in speech perception using probabilistic knowledge about speech production. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 244–263.

Ohala, J. J. (1975). Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasalfest: Papers from a symposium on nasals and nasalization* (pp. 289–316). Stanford: Language Universals Project.

Ohala, J. J. (1990). The phonetics and phonology of aspects of assimilation. In *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 258–275). Cambridge, UK: Cambridge University Press.

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven, & N. Warner (Eds.), *Laboratory phonology VII* (pp. 101–139). Berlin: Mouton de Gruyter.

Pitt, M. A. (2009). How are pronunciation variants of spoken words recognized? A test of generalization to newly learned words. *Journal of Memory and Language*, 61, 19–36.

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

Snoeren, N. D., Gaskell, M. G., & DiBetta, A. M. (2009). The perception of assimilation in newly learned novel words. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 35, 542–549.

Steriade, D. (2001). Directional asymmetries in place assimilation: A perceptual account. In E. Hume, & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219–250). New York, NJ: Academic Press.