The many ways listeners adapt to reductions in casual speech

Cover design: Tilman Harpe
ISBN: 978-90-76203-53-9
Printed and bound by Ipskamp Drukkers B.V., Nijmegen

The many ways listeners adapt to reductions in casual speech

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de Rector Magnificus, prof. mr. S.C.J.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen
op donderdag 12 december 2013
om 12.30 uur precies

door
Katja Monika Poellmann
geboren op 31 december 1982
te Marktredwitz, Duitsland

Promotoren:
Prof. dr. J. M. McQueen
Prof. dr. H. Mitterer (University of Malta)

Manuscriptcommissie:
Prof. dr. M. Ernestus
Prof. dr. A. Weber (Eberhard-Karls-Universität Tübingen)
Prof. dr. P. Zwitserlood (Westfälische Wilhems-Universität Münster)

The research reported in this thesis was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) under the priority program 1234 "Phonological and phonetic competence: Between grammar, signal processing, and neural activity". The work presented in Chapter 5 was funded by the European Science Foundation through a supporting grant received by NetWordS-09-RNP-089. Furthermore, this thesis was supported by the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

The many ways listeners adapt to reductions in casual speech

Doctoral thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus, Prof. dr. S.C.J.J. Kortmann,
according to the decision of the Council of Deans
to be defended in public
on Thursday, December 12, 2013
at 12:30 hours

by
Katja Monika Poellmann
born in Marktredwitz, Germany,
on December 31, 1982

Supervisors:
Prof. dr. J. M. McQueen
Prof. dr. H. Mitterer (University of Malta)

Manuscript Committee:
Prof. dr. M. Ernestus
Prof. dr. A. Weber (Eberhard-Karls-Universität Tübingen)
Prof. dr. P. Zwitserlood (Westfälische Wilhems-Universität Münster)

The research reported in this thesis was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) under the priority program 1234 "Phonological and phonetic competence: Between grammar, signal processing, and neural activity". The work presented in Chapter 5 was funded by the European Science Foundation through a supporting grant received by NetWordS-09-RNP-089. Furthermore, this thesis was supported by the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

Acknowledgements

During my first stay at the MPI, as an intern in the Adaptive Listening Group in 2009, I saw *The Island of Research* (http://profbailey.com/Seminar/island.htm) hanging next to my desk and was pretty impressed. Luckily, my Ph.D. time was not THAT hard and I did not have to go through all these challenging stages thanks to the help of various people. Unfortunately, I cannot mention all of them (and I would also be afraid of forgetting someone if I tried). So, knowing that there have been many more people who helped me on the way, I am especially grateful for the support of the following persons.

First of all, I would like to thank my supervisors, Holger Mitterer and James McQueen. Holger and James, you have been a fantastic team! Over the years, I have come to appreciate that more and more. I have realized that it should not be taken for granted that the two of you get along so well with each other and that your strong points complement each other so perfectly. Thanks to you, I did not end up in the Jungle of Authority. Holger, you were watching out for me, sending me the call for the NetWordS grant and forwarding me job ads for post-doc positions. James, you were super fast in commenting on my writings. Sometimes, I even got them back the very same day! Both of you thus made it possible for me to finish close to the end of my contract and helped me to avoid The Great Fundless Desert. I also managed to mostly steer clear of the Canyon of Despair. Sure, I ended up in the Delta of Editors with my first paper (Chapter 3), but your support at its first rejection was simply incredible! Holger, thanks a lot for your endless patience when I knocked on your door once again with yet another question. James, I always looked forward to getting your positive view on my results which, of course, hardly ever looked like I had wanted them to! Both of you contributed immensely to my Ocean of (Great) Experience and I could not have wished for better supervisors!

Next, I would like to thank Anne Cutler for providing such a stimulating research environment in the Language Comprehension Department. Anne, your never-ending enthusiasm for everything related to research is just infectious. You were never stingy with praise when talks at conferences and Fachbeirat went well. You also encouraged us Ph.D.

students to visit other research labs abroad to gain more experience. Thanks a lot for all of that!

Andrea, you are actually the one who opened the world (not only the island) of experimental and psycholinguistic research to me. You invited me for an internship back in 2009 and sent me to my first conference in 2010, from which resulted my first publication. Last year, you invited me to give a guest lecture in one of your courses and thus provided me with my first teaching experience. I am grateful for all the firsts that you have made possible for me and am happy that you agreed to be a member of my reading committee.

Of course, I am grateful as well to the other two members of my reading committee, Mirjam Ernestus and Pienie Zwitserlood. Mirjam and Pienie, thank you very much for taking the time (in a holiday period) to read my thesis and provide useful comments!

Over the years, I have taken part in the regular meetings of Anne's Comprehension Department (MPI), James' Sound Learning Group (BSI, Donders), and Mirjam's PI Group Speech Comprehension (CLS) and have had the opportunity to present my data there. I would therefore like to thank the (former) members of these groups for guiding me down the Peaks of Confusion (caused by my results) by offering suggestions, comments and ideas on my experiments.

In 2012, I spent four months in Harald Baayen's lab at the University of Tübingen to collect the data for Chapter 5. Well, to be more precise, I spent most of my time in the basement trying to get hold of participants' brain waves, but every time I got out of the Data Collection Den (mysteriously missing from the Island map), I enjoyed the easy and cheerful atmosphere of Harald's group. Harald, thanks a lot for the incredible amount of time you invested in analyzing the ERP data with me! I have learned a lot from my meetings with you and Jacolien and hope to learn even more in the future! Cyrus, Denis, and Fabian, without your help I would have been totally lost on the linux server when I was trying to get my models to run (and trying to avoid crashing the server…). I'll be eternally grateful for your assistance! Furthermore, I would like to thank Sam Featherston and Carolin Dudschig, who helped me to get the EEG lab running in the first place. Beate Starke deserves a very special thank you for all her personal input, for speedily ordering the equipment I asked for, for providing me with all the small things like towels, gel, a pin board, a stool, etc. that I needed to run the lab smoothly. Beate, was hätte ich nur ohne dich getan? Vielen herzlichen Dank für deine Hilfe!!! When my four months in Tübingen were drawing to a close and I was getting more and more desperate to find the last Northern German participants that I still needed,

Anna Pryslopska came to my rescue. Anna, thanks so much for sending your participants to my lab as well, thereby helping me find my final subjects!

In Nijmegen it was equally difficult to find enough right-handed Northern Germans (especially males!) willing to participate in my EEG experiment. I would therefore like to thank all the German IMPRS students who agreed to endure my experiment for chocolate only! Thanks also to all the people who sent their German friends, roommates, colleagues and students my way. Thordis, thank you for motivating your HAN students to take part in my experiment! Kimberley, thanks for all the names you suggested to me. Atsuko, thank you so much for driving German male PhD students and post-docs from the Donders into my waiting hands!

Whenever I struggled with the smooth run of experiments in Nijmegen, I could count on the help of the TG. Alex and Johan, thanks a lot for getting me out of the Jungle of Recalcitrant Devices (which has also mysteriously evaded discovery by the cartographers of the Island!). Loes en José, thank you for helping me interact with my Dutch participants more easily by practicing Dutch with me!

Marijt, you were the voice of most of my eye-tracking experiments. While Holger was totally thrilled that you produce /b/-reductions naturally, I sometimes had difficulties to find a clear /b/ in your recordings—and then had to ask you for more examples of your speech ;-)) Thanks a lot for being so patient with me! It was very funny listening to your slips of the tongue again later on, and I'm sure that Patrick and Jiyoun must have wondered occasionally what it was again that I was laughing about.

Eva, you agreed, and probably regretted later on, to be the voice for my EEG experiment. I really was not aware of the torture I put you through with my at times long and complicated filler sentences. Thanks, not only for your help with the recordings, but also for your invaluable praat scripts. It was nice meeting you again at conferences and your advice concerning job search, networking, and academic life in general was very helpful. Thanks for all your help, Eva!

What are friends and colleagues for, if not for preventing you from ending up totally exhausted in the Forest of Fatigue? I enjoyed many a refreshing chat in the corridor, on the new couch, during pilot studies (with me as a participant) and EEG preparations (with me as the experimenter). Thanks to all my colleagues who felt like taking a small break with me! Of course, the more drinks (e.g., hot chocolate!!) and/or sweets and/or food (in the best case: all of them!) accompanied the conversations, the nicer it was. I thus have fond memories not only of solving work & life problems and feeding birds in winter time, but also of cooking dinner with my officemates Patrick and Jiyoun. Jiyoun, your happy personality and big smile helped me overcome my listlessness so many times. My Ph.D. time would certainly have been a lot less fun without you! Patrick, thanks a lot for all your advice and tips concerning Photoshop, EndNote, the creation of a nice manuscript for the printer and arrangements for the defense. I tried my best to benefit from your experiences! Wencui, you joined our office only later, but thanks a lot for the Chinese and other sweets you shared so generously with us! Carmen, you spent such a huge amount of your free time showing Eva and me around when we were at the ICPhS in Hong Kong. Basically every night, you took us to another fantastic restaurant. I was so touched and am still amazed by your caring hospitality. Sarah, thanks for some reassuring chats over nice hot chocolate! Laurence, I have no clue how German my Dutch summary would have been without you! Thanks a lot for your corrections and help! Neil, thank you so much for the piles of information on Boston you provided me with! Thordis, I was always happy when I was invited to come to one of your parties across the border! You also provided me with delicious Lindt chocolate in times of great need (e.g., during the seemingly endless formatting period of my thesis). Salomi, I will miss baking Christmas Plätzchen with Jiyoun and you and even more going out for pannenkoeken with you. I was so touched that you wanted to come with me to the eye specialist when I had to have my eyes lasered! You are a fantastic and considerate friend who cares so much for others! Thanks a lot for that! I am really happy that you and Thordis will be at my side at my defense! Salomi and Thordis, thanks a lot for agreeing to be my paranymphs!

I visited the Bay of Leisure not only with my Nijmegen friends, but also with my friends back in Germany. The latter trips may have been rarer but they were then also more intense. Steffi und Isa, ich habe mich super gefreut, dass ihr beide mich hier in Nijmegen besucht habt! Vielen lieben Dank, dass ihr mich nicht vergessen habt! Ich freu mich immer auf nen Spieleabend bzw. nen laaangen Kaffee mit euch! Gerade Letzteres gilt auch für dich, Thomas! Saskia, Julia und Jeka, ich freu mich schon jetzt auf unseren nächsten Mädelstreff samt intensiver Gespräche und bin gespannt, wo es uns dann hinverschlägt! Liebe Saskia, vielen lieben Dank für die langen Telefonate! Ich konnte immer auf dich zählen, wenn ich mal jemanden zum Reden brauchte! Vielen lieben Dank, dass ihr alle mich auch in der Ferne über Häuserbau, Hochzeiten, Schwangerschaften und Ähnlichem auf dem Laufenden gehalten habt!

I also owe more than a HUGE thank you to my family. Over the years, my aunt and uncle have visited me wherever I went abroad—France, Italy, the Netherlands. And whenever I visited them, they provided me with a second home. Karin und Hans, vielen lieben Dank für's Bekochen, Bemuttern und Verhätscheln! Kein Wunder, dass ich mich bei euch so zu Hause fühle! My brother also visited me and kept me updated via skype. Michi, deine (Alltags-)Erlebnisse haben mich oft so zum Lachen gebracht, dass ich schon beinah geheult habe! Ich hoffe, dass wir noch viel mehr witzige Gespräche (hoffentlich immer auf deine Kosten!) haben werden. Of course, I would not have been able to achieve all the things I have without the help, support and love of my parents. Mutter und Vater, ihr habt mich immer und in jeder Hinsicht unterstützt und mich ermutigt, meinen Weg zu gehen. Vielen, vielen Dank für all die Freiheiten, die ihr mir ermöglicht habt!

Finally, my boyfriend deserves a BIIIIIIG thank you as well. Remon, you have shared my time in the Netherlands with me and have surely challenged me in various ways—which has helped me to grow stronger. You have definitely made my life more exciting, but you have also helped me to relax and enjoy life. Importantly, you have also supported me in many ways (e.g., when moving, applying for jobs and making decisions about my future career). Thank you for all of that! I hope that we will share many more fantastic moments together!

Table of contents

Introduction

Chapter 1

In everyday life, we hardly ever have conversations in ideal listening situations. There is always at least some background noise present in our environment. Speech recognition may also be hampered by the listener's hearing ability or the speaker's way of producing speech (see Mattys, Davis, Bradlow, & Scott, 2012, for a review on speech recognition in adverse conditions). Problems due to speech production are obvious when one is talking to a non-native speaker who has a strong foreign accent or to a native speaker who speaks an unfamiliar dialect. The speaker's pronunciation habits can also pose problems for the listener if the speaker talks in the standard variety, but does not pay much attention to articulating speech sounds carefully. It is indeed a very common phenomenon that speakers weaken and delete segments and entire syllables when they speak casually (Ernestus, 2000; Ernestus & Warner, 2011a; Johnson, 2004). For Dutch, Schuppler, Ernestus, Scharenborg, and Boves (2011) found that in 31.2% of all word tokens in the Ernestus Corpus of Spontaneous Dutch (Ernestus, 2000), segments were changed or missing. Syllables were reduced in 9.1% of all polysyllabic word tokens in the corpus. A Dutch speaker may for instance pronounce the word-initial /b/ in *banaan* 'banana' as [m] and say *manaan,* or s/he may shorten the prefix *ver-* to [f] and say *floren* instead of *verloren* 'lost'. In the case of *manaan,* the word-initial voiced stop is weakened as it becomes a nasal and therefore less consonant-like (see definition of "lenition" in Trask, 2000).

These and similar kinds of deviations from canonical pronunciations have attracted a lot of attention in recent years, leading to the publication of several special issues in different journals (Ernestus & Warner, 2011b; Mattys, Bradlow, Davis, & Scott, 2012; Mitterer, 2012). In this broad research area, two themes have been developed independently: research on reductions and research on speaker adaptation.

Reductions are optional. The pronunciation habits of the speaker thus play a crucial role in determining whether or not reductions occur. There are of course contexts in which they are more likely to occur, but nothing restrains the speaker from articulating words in their full form (Van Son & Pols, 1990, 1992). Furthermore, when speakers reduce words, they can do so in many different ways. Instead of reducing the /b/ in /b/-initial words to [m], the same speaker could just as well reduce it to the labio-dental approximant [ʋ] and pronounce *banaan* as *wanaan*. As speakers usually have way more than one option to reduce a given word, they vary in the choices they make (Ernestus, Baayen, & Schreuder, 2002). In addition, Keune, Ernestus, van Hout, and Baayen (2005), who investigated how common reductions are among different talker groups (male and female speakers of Dutch and Flemish), found that Dutch men reduced most often, while Flemish highly educated women were least likely to do so. Reductions thus depend on the speaker in several ways. On the perception side, however, there has been little research on these talker differences in reduction style.

Research on speaker adaptation in other domains, in contrast, has shown that native listeners can adapt rapidly to regional or foreign accented speech (e.g., Clarke & Garrett, 2004; Floccia, Goslin, Girard, & Konopczynski, 2006; Witteman, 2013). If they have been exposed to multiple non-native speakers, native listeners can also take advantage of their knowledge of the accent to understand a novel non-native speaker of the same accent better (Bradlow & Bent, 2008). Furthermore, it has been shown that non-native listeners can adapt more easily to (regional accents of) a foreign language, if they watch movies with foreign subtitles (Mitterer & McQueen, 2009a).

As closely related as research on reductions and speaker adaptation may seem to be, there was, at the beginning of this endeavor, no work combining these two lines of research. That is why this approach was used in this thesis: It has become evident that, on the one hand, speakers differ in their likelihood and in their way to reduce speech and, on the other hand, listeners are able to adapt to speakers whose pronunciations deviate to a large extent from standard pronunciations. Native listeners may thus very well be able to adapt to an individual speaker's tendency to reduce. The experiments in this thesis will ask if this is the case.

Both research on reductions and research on adaptation (or perceptual learning) have also fueled a major theoretical debate in the field: How are words stored in the mental lexicon and consequently, how are spoken words recognized? There are basically two extreme answers to these questions and many intermediate (hybrid) possibilities. One view, advocated by exemplar-based models of spoken-word recognition, is that the speech signal is mapped directly onto a lexicon that consists of acoustically-detailed representations. The other view, advocated by abstractionist models, claims that the lexicon consists of phonologically abstract representations. The speech signal, therefore, has to be mapped first onto abstract intermediate (prelexical) representations (e.g., phonemes) before the lexicon can be accessed. Certainly, a strongly reduced word like Dutch *mok* for *mogelijk* 'possible' can be more efficiently recognized if it has a separate entry in the mental lexicon than by applying multiple complex mapping rules—if it is even possible to generate such rules. Abstraction processes are beneficial, however, if segments deviate in a regular manner from the canonical pronunciation, for example, when an American English dialect speaker raises the vowel /æ/ before /g/, but not before /k/. Through abstract mapping rules, learning can be applied to new words that show the same pattern (/g/-words) and, in this case, facilitate the recognition of words that did not undergo the dialect shift (/k/-words; see Dahan, Drucker, & Scarborough, 2008).

In this thesis, I will investigate whether listeners can adapt to segmental and syllabic reductions, that is, to words that are reduced in a regular way. This research will shed light on the role of storage and abstraction processes for recognizing regularly reduced forms and hence contribute to the debate on models of spoken word recognition. Furthermore, investigating the time course of the learning process and other possible constraints on learning will be major goals.

Previous research on reductions

Factors such as frequency of occurrence and contextual predictability, as well as stress pattern, play an important role in determining whether or not reductions will occur. Usually, the more frequent or predictable certain linguistic units (e.g., segments, syllables, words) are, the more likely it is that they undergo reduction processes (Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Pluymaekers, Ernestus, & Baayen, 2005). Furthermore, the phonological context has to be adequate. In general, reductions occur typically in unstressed syllables. To give a more specific example, stops are frequently reduced to approximants if they occur in unstressed intervocalic position (Warner & Tucker, 2011).

Research on the comprehension of reduced word forms has shown that listeners are sensitive to these factors and use the sentence and wider discourse context to recognize reduced words (Brouwer, Mitterer, & Huettig, 2013; Ernestus, et al., 2002). Strongly reduced words cannot even be understood out of linguistic context (Ernestus, et al., 2002). Mitterer and McQueen (2009b) showed that listeners also use probabilistic cues from the surrounding context to compensate for reductions: If word-final /t/ is preceded by an /s/ and followed by a bilabial sound (/b/, /m/), the /t/ is very likely to be reduced in Dutch. Dutch listeners were found to be sensitive to these probabilistic facts when compensating for /t/-reductions. Furthermore, listeners take lexical knowledge into account. They are, for instance, more likely to restore a reduced /t/ if that gives rise to an existing word (Mitterer & Ernestus, 2006). Listeners can also make use of the tendency that high-frequency words are more likely than low-frequency words to be reduced. They assume, for instance, that a reduced prefix is more likely to belong to a high-frequency word than to a low-frequency word (Mitterer & Russell, 2013). In addition, listeners exploit fine phonetic detail in order to recover from reduced forms. Manuel (1992) reported that listeners are able to distinguish a reduced form [spɔːt] of ‘support’ from the unreduced form [spɔːt] when it meant ‘sport’.

Listeners thus have different ways of dealing with reduced speech at their disposal and use information from a wide variety of sources, such as sentence context, the lexicon and fine phonetic detail. Apparently, listeners use these mechanisms unconsciously, as they are usually not aware of missing segments in reduced words (Kemps, Ernestus, Schreuder, & Baayen, 2004).

In this thesis, I examine whether listeners, in addition to the mechanisms just mentioned, also use adaptation processes in order to better recognize (regularly) reduced word forms. While previous research on the comprehension of reduced forms has concentrated on whether listeners use cues that depend on the language (e.g., probabilistic factors such as word frequency) and/or on the speech signal (e.g., fine phonetic detail) in order to deal with reductions, this thesis tests whether listeners also use cues that depend on the speaker by testing whether they can adapt to reduction styles that are typical for a certain speaker through exposure to that speaker talking in that specific style, relative to exposure to the same speaker talking in a different way.

Perceptual learning

Research on perceptual learning about speaker idiosyncrasies—an adaptation process that is potentially very similar to adaptation to segmental reductions—has provided evidence that words are recognized through intermediate units of perception. Listeners were able to generalize learning about an odd pronunciation of a speech sound (imagine, for example, a speaker that lisps) to other words containing that sound (McQueen, Cutler, & Norris, 2006). In this and similar perceptual learning studies (for a review, see Samuel & Kraljic, 2009), one group of participants was trained, based on lexical bias, to perceive an ambiguous sound [$^{s}/_{f}$] as /s/, and another group was trained to perceive the same sound as /f/. That is, the first group heard the ambiguous sound in words like English *horse* ("hor[$^{s}/_{f}$]"), while the second heard it in words like English *giraffe* ("gira[$^{s}/_{f}$]"). As the ambiguous sound occurred late in the word, sufficient lexical knowledge was available to guide the retuning.

Such perceptual learning effects have been shown to arise fast (Kraljic & Samuel, 2007) and are claimed to reflect automatic, prelexical processes (McQueen, Cutler, et al., 2006; McQueen, Norris, & Cutler, 2006). Listeners did not have to make explicit judgments about the ambiguous sound for learning to take place (McQueen, Norris, et al., 2006) and even just listening to a story that contained ambiguous sounds induced a learning effect (Eisner & McQueen, 2006). Moreover, these learning effects proved to be long-lasting (Eisner & McQueen, 2006). However, using a lexical decision task in the exposure phase and therefore presenting words in isolation, Jesse and McQueen (2011) did not observe a perceptual learning effect for words that contained the ambiguous sound in word-initial position. They concluded therefore that lexical knowledge has to be available when the ambiguous sound is being processed. Furthermore, Kraljic, Samuel, and Brennan (2008) showed that listeners have to interpret the ambiguous sound as characteristic of the speaker's pronunciation. If they can attribute the odd sound to incidental factors (e.g., a speaker holding a pen in his or her mouth) no perceptual learning effect was found.

While the studies described so far tested short-term learning effects, Sumner and Samuel (2009) investigated a long-term learning effect by examining the effect of experience on the comprehension of dialect variants. They tested three groups varying in their experience with r-dropping in *–er* final words that is common in the New York City area. While one group was unfamiliar with r-less dialects, another group actively produced r-less variants, whereas an intermediate group was familiar with the variants but did not produce them. The r-less variants posed problems for the inexperienced group, whereas the two experienced groups had no difficulties understanding them. Both showed short-term priming effects for these words. However, only the group who actively produced the r-less variant forms also showed long-term priming effects for them. Sumner and Samuel took this as evidence that only the group who produced the r-less variants had stored them as such in the mental lexicon, while they argued that participants in the intermediate group are "fluent listeners" who exhibit immediate perceptual abilities to recognize r-less word forms.

Adaptation processes not only have been studied for speakers that pronounce a certain phoneme in an idiosyncratic way or for speakers that have a regional accent, but also for speakers with a foreign accent. Reinisch and Weber (2012), for instance, showed that listeners can adapt quickly to a non-native stress pattern. Furthermore, familiarity with the foreign accent is beneficial and facilitates the recognition of strongly accented words, for instance words in which a non-native speaker substitutes a target (L2) vowel (e.g., Dutch /œy/) by one of his or her L1 (e.g., German /ɔɪ/) that differs considerably from the L2 vowel (Witteman, Weber, & McQueen, 2013).

This thesis tests whether similar kinds of long-term and short-term learning effects can be found for segmental and syllabic reductions and whether familiarity with reduction types is necessary for learning to occur.

Models of spoken word recognition

The aim of this thesis is to investigate what kind of learning mechanisms listeners use when dealing with reduced forms. Listeners might store reduced variants in the mental lexicon and/or they may build up abstract mapping rules to compensate for regular reduction processes and apply this knowledge to previously unheard reduced words. Finding evidence of these kinds of processes will contribute to the debate on models of spoken word recognition.

These models can be classified on a continuum ranging from purely episodic to purely abstractionist. They differ in their assumptions on the kind of information that is stored in the mental lexicon and therefore also on whether lexical access takes place directly or only after an intermediate (prelexical) stage. In extreme episodic models, every instance of all words that the listener has ever encountered is preserved in all its fine phonetic detail, including characteristics of the speaker, speech rate, etc. (e.g., Goldinger, 1998; Johnson, 1997). Thus, an entry in the mental lexicon not only consists of various memories of a word's citation form, but also contains episodic traces of reduced forms showing different kinds and/or degrees of reduction. In episodic models, the incoming speech signal is compared directly to all exemplars stored in the mental lexicon. Extreme episodic models can thus easily account for frequency effects and for the recognition of severely reduced word forms such as [mok] for Dutch /moxələk/ *mogelijk* 'possible', but have problems making generalizations about sublexical units (e.g., segments, syllables) and applying these to new words (Cutler, Eisner, McQueen, & Norris, 2010).

At the other endpoint of the continuum, purely abstractionist models assume a sparse lexicon that consists of abstract representations (e.g., strings of phonemes) of a word's canonical pronunciation. The acoustic input is therefore mapped onto abstract phonological units, which may be features (Gaskell & Marslen-Wilson, 1997), phonemes (Norris & McQueen, 2008) or features and phonemes (McClelland & Elman, 1986) before the lexicon is accessed. Prelexical mapping rules have thus to compensate for continuous speech phenomena such as reductions. These mapping rules are generated based on a small sample of exemplars which are then discarded. Purely abstractionist models have difficulties handling irregular variation that only affects a small number of words (e.g., severe reductions in very high-frequency words), but can easily account for generalization effects of regularly occurring variation across words (McQueen, Cutler, et al., 2006).

In between these two endpoints of the continuum lie hybrid models, models that contain both a redundant lexicon and abstract generalizations to a varying degree. McLennan, Luce, and Charles-Luce (2003), for instance, proposed a hybrid model that consists of abstract chunks (i.e., representations for words, allophones, phonological features) as well as exemplar-based chunks (e.g., speaker information). These chunks are activated by and resonate with the acoustic input which leads to the listener's percept.

The model proposed by Ranbom and Connine (2007) preserves phonological variation but abstracts away from other sources of variation such as speaker information, speech rate, etc. According to this account, the mental lexicon consists of multiple abstract phonological representations for individual words, with stronger representations for the most frequently occurring forms. This model does not require special processes to compensate for variation such as reductions in the speech signal.

Methodology

This thesis examines adaptation to reductions in laboratory settings. That is, I use speech corpora only to verify that the reduction types under investigation occur frequently enough in spontaneous conversations, but I do not use extracts of these corpora in my experiments. The latter has been done by Brouwer (2010), who investigated how strongly reduced words are processed compared to canonically pronounced words. She found, for instance, that listeners adjust to reduced speech and penalize acoustic mismatches less strongly than when hearing laboratory speech. Furthermore, she reported that strongly supportive discourse context is more beneficial for recognizing reduced words than for recognizing unreduced words in natural, communicative settings. An approach using corpus speech is, however, less suitable for testing learning effects. One reason for this is that the speech of some corpora was not recorded under ideal conditions so that it contains a considerable amount of background noise. Another point is that using corpus speech constrains the selection of stimuli. There might just not be sufficient word types that show reduction and that are reduced in a comparable way to run a learning experiment. Moreover, the context in which the reduced words occur cannot easily be controlled for. It might thus be difficult to draw any conclusions about adaptation processes if it cannot be ruled out that better recognition of reduced words was due to supportive factors in the surrounding context. Instead of extracting real-world reductions from corpora, I therefore recorded speakers and manipulated their recordings to imitate naturally occurring reductions. In this way, I could model a speaker who reduces consistently and I had better control over the stimuli.

My basic question is whether listeners can benefit from exposure to reductions to better recognize other reduced forms. Furthermore, I am interested in the time course of perceptual learning about reductions. This raises the question how one can measure the efficiency of word recognition online.

The visual-world paradigm

The efficiency of the word recognition process can be measured online using the visual-world paradigm. In Chapters 2-4, I use printed-word variants (McQueen &

Viebahn, 2007; Mitterer & McQueen, 2009b) of the classical eye-tracking paradigm (Allopenna, Magnuson, & Tanenhaus, 1998) to investigate lexical activation as reduced words are being processed.

In an eye-tracking experiment, participants see a number of pictures or printed words on a computer screen. While they listen to sentences (and follow instructions), their eye movements are monitored. The reasoning behind this is that we look unconsciously to objects (in the real world and on pictures) and printed words as soon as we hear them mentioned. The number of looks to a target or a competitor indicates how strongly those words are activated.

In the studies presented in this thesis, there will be the orthographic form of the reduced target word, one phonological competitor and two unrelated distractors on the screen. The competitor overlaps in the onset with the reduced form of the target word (e.g., Dutch *waarom* 'why' is a competitor to the reduced form *waron* of *baron* 'baron').

In the classical eye-tracking paradigm (Allopenna, et al., 1998), pictures instead of printed words were presented on the screen. The advantage of using printed words is that the experimenter has more freedom in choosing the experimental items, as it is possible to also include items which cannot be depicted easily.

The difference in target activation relative to competitor activation will tell us how efficient listeners are in recognizing reduced target words. The bigger the difference (i.e., the more the reduced target and the less the canonical competitor is activated) the easier it is for listeners to recognize the reduced target word. Observing a greater target preference (over a competitor preference) for a group of listeners that is experienced with a given reduction type compared to one or more groups that have not encountered this reduction type before will be taken as evidence of adaptation.

Event-related brain potentials (ERPs)

Another way to measure the efficiency of word recognition online is by using the event-related potential (ERP) technique. This technique is used in Chapter 5. Event-related potentials have mostly been used to show when words are *difficult* to recognize, by, for instance, putting them in unlikely contexts. The classic example here is: *"I take a coffee with cream and* ***dog****"*. The sentence ending *dog* will elicit a strong negative-going deviation around 400 ms after word onset (an N400 component) compared to the expected sentence ending *sugar*. N400 effects have also been observed when participants had to process nonwords (for a review on the N400, see Kutas & Federmeier, 2011). In this thesis, I ask whether similar effects can be found for the processing of reduced words compared to unreduced words. If listeners do not recognize reduced words, these might be processed similarly to non-existing words and elicit an N400 effect. In addition, listeners might not expect the speaker to leave out certain sounds of a word. Therefore, an N200 might be observed. This is a component that has been shown to reflect detection of a phonological mismatch (Connolly & Phillips, 1994; Van den Brink, Brown, & Hagoort, 2001). The amplitude of an N200 or N400 component is taken as evidence of processing difficulty (the greater the amplitude the more processing involved). Therefore, adaptation to reduced forms will be visible if ERP effects get smaller over the course of the experiment. Listeners then have less difficulty recognizing the reduced words.

Compared to tracking eye-movements, ERPs are a more direct measure of spoken word processing because no pictures or printed words have to be presented to listeners. Furthermore, no strong (real-word) competitors have to be found. This provides more freedom in choosing critical stimuli and allows also for reductions that lead to phonotactically illegal sequences to be investigated.

Outline of the thesis

The major objective of this thesis is to investigate whether and how native listeners adapt to regularly reduced speech and what the time course of this learning process looks like. To that end, online methods such as eye-tracking and ERPs will be used. Adaptation to reductions of different sizes (i.e., segmental and syllabic reductions) will be investigated and possible constraints on this learning process such as familiarity with a given reduction type will be explored. The results will inform us what mechanisms and representations are necessary for recognizing reduced words that deviate in varying degrees from canonical pronunciations.

Chapter 2 focuses on the time course of perceptual learning about speaker idiosyncrasies (here, an odd pronunciation of a fricative between /s/ and /f/) and tests how many items are necessary for learning to occur. Another aim is to examine, using the eye-tracking paradigm, whether learning emerges gradually or in a step-wise fashion. Kraljic and Samuel (2007) reported that 10 ambiguous tokens were enough to induce a stable learning effect. Chapter 2 tests whether eye-tracking can provide deeper insights into a well-established learning process about speaker idiosyncrasies which are likely to be comparable to certain segmental reductions—reductions in which the articulation of a segment is weakened so that it results in an ambiguous sound.

Using also an eye-tracking paradigm, Chapter 3 investigates whether (Dutch) listeners can adapt to regular segmental (/b/ → [ʋ]) and syllabic (*ver*- → [fː]) reductions in the short term, that is, within an experimental session. A segmental reduction group will be exposed to reduced /b/-words and a syllabic reduction group to reduced *ver*-words. In the test phase, both groups will hear both types of reduction applied to new words. Chapter 3 tests thus whether there is generalization of learning about reductions to new words. In the case of *ver*-reductions, this would have consequences for abstractionist models of spoken word recognition, which do not include prelexical representations larger than a segment. The latter would be required to account for generalized learning about *ver*-reductions.

While Chapter 3 examines adaptation to the reduction of the prefix *ver*-, Chapter 4 asks whether listeners can adapt to reduced syllables that are not morphemes. To that end, a syllabic reduction group will be exposed to words in which a full vowel is deleted. Furthermore, a segmental reduction group will be exposed to another variant of reduced /b/-words (where the word-initial /b/ is reduced to [m]). Apart from generalization of learning *within* a given reduction type, Chapter 4 explores further adaptation processes, such as generalization of learning *across* reduction types and word-specific learning. To investigate the former, an additional control group will be tested who will not receive any exposure to reductions. In the test phase, all three groups will hear both reduction types, while their eye movements are being tracked. In Experiments 4.1 and 4.2, the reduced words of the exposure phase are repeated in the test phase. In Experiment 4.3, the reduced words in the test phase will not have appeared in the exposure phase. While Experiments 4.1 and 4.2 test word-specific learning, Experiment 4.3 thus tests generalization of learning within reduction types. All three experiments test generalization of learning across reduction types. Chapter 4 will thus provide an overview of possible adaptation processes involved in recognizing reduced word forms.

The last experimental chapter investigates adaptation on a longer time frame than the previous experiments. While Chapters 3 and 4 look at the effect of 10 minutes of exposure on performance in a subsequent test phase, Chapter 5 uses a quasi-experimental approach. Participants who vary in their experience with a given type of reduction are tested. This can be achieved by using patterns of reductions that are regionally constrained, specifically, the reduction of the prefix *ge*- to [g]/[k] common in Southern German and Austrian German. Participants will also be confronted with another type of prefix reduction that is relatively unfamiliar throughout Germany and Austria (reduction of the prefix *ver*- to [f]). The ERPs of these reduced forms in Northern Germans, migrant Northern Germans living in the South and Southern Germans will be compared to those of control participants who hear unreduced versions of the same words. Chapter 5 thus explores differences in processing of and adaptation to reduced forms due to prior familiarity with these reduction types. Moreover, the processing of reductions leading to illegal phonotactic sequences compared to that leading to legal phonotactic sequences will be examined.

Finally, Chapter 6 summarizes the results and provides a general discussion of the main findings of this thesis.

The time course of perceptual learning

Chapter 2

Based on Poellmann, K., McQueen, J. M., & Mitterer, H. (2011). The time course of perceptual learning. In W.-S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011 [ICPhS XVII]* (p. 1618-1621). Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong.

Abstract

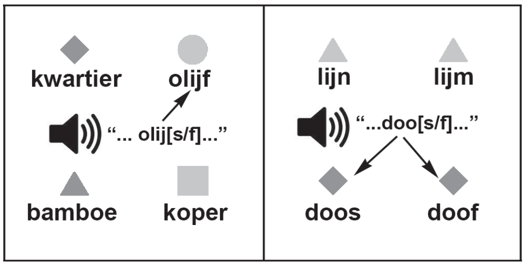
Two groups of participants were trained to perceive an ambiguous sound [$^{s}/_{f}$] as either /s/ or /f/ based on lexical bias: One group heard the ambiguous fricative in /s/-final words, the other in /f/-final words. This kind of exposure leads to a recalibration of the /s/-/f/ contrast (e.g., Norris, McQueen, & Cutler, 2003). In order to investigate when and how this recalibration emerges, test trials were interspersed among training and filler trials. The learning effect needed at least 10 clear training items to arise. Its emergence seemed to occur in a rather step-wise fashion. Learning did not improve much after it first appeared. It is likely, however, that the early test trials attracted participants' attention and therefore may have interfered with the learning process.

Introduction

Listeners can adapt to an odd pronunciation of a given phoneme. More precisely, listeners learn to interpret an ambiguous sound between /s/ and /f/ ([$^{s}/_{f}$]) as either /s/ or /f/ according to the training they received: Participants who heard [$^{s}/_{f}$] in words like "gira[$^{s}/_{f}$]" learned to interpret the sound as [f], while participants who heard [$^{s}/_{f}$] in words like "platypu[$^{s}/_{f}$]" learned to interpret it as [s] (Eisner & McQueen, 2005; Norris, et al., 2003; Sjerps & McQueen, 2010). So far, all perceptual learning studies have used separate exposure and test phases. In the exposure phase, participants usually made lexical decisions on (non-) words they heard (Eisner & McQueen, 2005; Kraljic & Samuel, 2007; Norris, et al., 2003; Sjerps & McQueen, 2010). In the test phase, they categorized a range of ambiguous sounds (Eisner & McQueen, 2005; Kraljic & Samuel, 2007; Norris, et al., 2003) or did a cross-modal priming task (Sjerps & McQueen, 2010).

Exposure to just 10 instances of an ambiguous sound has been found to be enough to produce a stable perceptual learning effect (Kraljic & Samuel, 2007). However, the division of the experimental session into separate exposure and test phases has impeded a closer look at how learning emerges over time. The current study therefore used a variant of the visual-world eye-tracking paradigm (Mitterer & McQueen, 2009) to examine the time course of perceptual learning. Participants saw displays as in Figure 1 and heard an instruction to click on one of the four words on the screen. Such trials can be learning trials when an ambiguous fricative appears in an unambiguous position (as in the left panel of Figure 1) or test trials when the ambiguous fricative appears at the end of a minimal pair (as in the right panel of Figure 1). This allows us to ask two questions: First, how many critical items are necessary to trigger perceptual learning? Second, is the learning process gradual or does it emerge suddenly in a step-wise fashion after a certain number of exposure trials?

Figure 1. Example of printed-word displays for a training trial (left panel) and a test trial (right panel). When hearing an [$^{s}/_{f}$]-bearing test word (e.g., *doo[$^{s}/_{f}$]*), participants should look to the /s/- or /f/-word according to their training. Participants saw only the four words and the four shapes on the screen.



Method

Participants

All participants were paid native speakers of Dutch. Twelve took part in each of two pretests and 44 in the main eye-tracking experiment.

Stimuli

Stimulus selection. We selected three types of stimuli containing [s] or [f] which could be replaced by the ambiguous fricative $[^{s}/_{f}]$: 20 training items, in which the critical fricative could only be interpreted as [s] or [f] (e.g., *radijs* ‘radish’ and *olijf* ‘olive’; *radijf* and *olijs* are nonwords in Dutch), 20 temporary minimal pairs, in which the ambiguous fricative could temporarily be interpreted as either [s] or [f], but was disambiguated later in the word (e.g., *gister* ‘yesterday’ - *giftig* ‘toxic’), and 20 minimal pairs (e.g., *doos* ‘box’ - *doof* ‘deaf’), which constituted the test items. The temporary minimal pairs were supposed to serve as test and training items at the same time. The corresponding sets of stimuli were matched in terms of frequency (see Table 1). When pronounced with a natural fricative, the training items of one group served as contrast items for the other group. Thus, depending on the condition (/s/-bias or /f/-bias), it varied as to which words participants heard with natural or ambiguous fricatives (see Table 2).

Table 1. Mean frequencies per million of the different types of stimuli.

Items	/s/-words	/f/-words
training	9.9	9.5
test pairs	31.2	35.5
temporary minimal pairs	5.7	7.7

Table 2. Examples of the different stimulus types (natural and $[^{s}/_{f}]$-bearing) for both groups.

Items	/s/-bias group	/f/-bias group
training	radij$[^{s}/_{f}]$	olij$[^{s}/_{f}]$
contrast	olijf	radijs
test pairs	doo$[^{s}/_{f}]$ + doof	doo$[^{s}/_{f}]$ + doos
temporary minimal pairs	gi$[^{s}/_{f}]$ter	gi$[^{s}/_{f}]$tig

Stimulus construction. Digital recordings of the stimuli were made by a male native speaker of Dutch in a sound-proof booth, sampling at 44 kHz. All target items containing /s/ or /f/ were recorded in their natural version. The /f/-targets were additionally recorded with an

[s] replacing the [f] (e.g., *olijf* 'olive' as well as *olijs*). When creating the ambiguous $[^{s}/_{f}]$-items, the [s]-versions were used: The [s] was removed and replaced with instances of $[^{s}/_{f}]$, which were selected by means of two pretests. To reduce a coarticulatory [s]-bias, splicing points were chosen reasonably early in the vowel preceding the fricative. In addition, the amplitude of that vowel was decreased to match better an ambiguous sound that could still be interpreted as [f] by a native speaker of Dutch.

In order to find ambiguous fricatives midway between [s] and [f], two pretests were run. For the first pretest, 14 continua of VC-syllables were created. These contained all the vowels present in the 60 critical stimuli and a digital mixture of [s] and [f] varying in the proportions of the two fricatives in 11 steps (from 100% [s] 0% [f] via 90% [s] 10% [f] to 0% [s] 100% [f]). Participants were asked to categorize the fricatives of every second step (i.e., five steps per syllable) as either [s] or [f]. For each of the 14 vowels, the step whose value was closest to the ideal value of 50% of [s] responses was chosen as the ambiguous fricative for the second pretest.

As in the main experiment, the second pretest used $[^{s}/_{f}]$-bearing words embedded in complete sentences. Participants heard the carrier sentence “Klik op het woordje” (‘Click on the word’) followed by a member of an /s/-/f/-minimal pair (e.g., *doof*). These stimuli had either a natural [f], a natural [s], or an ambiguous fricative $[^{s}/_{f}]$ (which varied according to the preceding vowel). Every participant heard and categorized each item four times. Ambiguous fricatives were selected per vowel context if they were categorized as [s] between 30% and 70%. For those outside this range, a new step was chosen in order to correct for the observed /f/- or /s/-bias.

Design and Procedure

Each participant heard 240 instructions like “Klik op het woordje *doo*$[^{s}/_{f}]$ onder het rooie ruitje” (‘Click on the word *doo*$[^{s}/_{f}]$ underneath the red diamond’, see Figure 1) and was asked to follow these instructions using the computer mouse. Except for the test trials, there were always four different shapes (a circle, a diamond, a triangle, and a rectangle) on the screen. The two shapes associated with the target and its same-color competitor were presented in one of four colors (red, green, yellow and blue), the shapes paired with the two distractors were presented in one of the three remaining colors (see left panel in Figure 1 with gray-scale coding of different colors). In the test trials, the target and competitor shapes and those for the two distractors were the same. That is, there were only two different shapes on the screen (see right panel in Figure 1). This design was chosen in order to force participants to make a decision about every word they heard.

Eye movements were tracked using an Eyelink 1000 at a sampling rate of 1 kHz. The stimuli were presented in 20 miniblocks, each consisting of one training item, one contrast item, both members of a minimal pair (an [$^{s}/_{f}$]-bearing one serving as test item, the other one with a natural fricative serving as a second contrast item), one member of a temporary minimal pair (see Table 2), and seven filler trials. In order to be able to interpret the observed effects correctly, we used only one randomization of the 240 trials for all participants.

Results

Mouse-click Responses

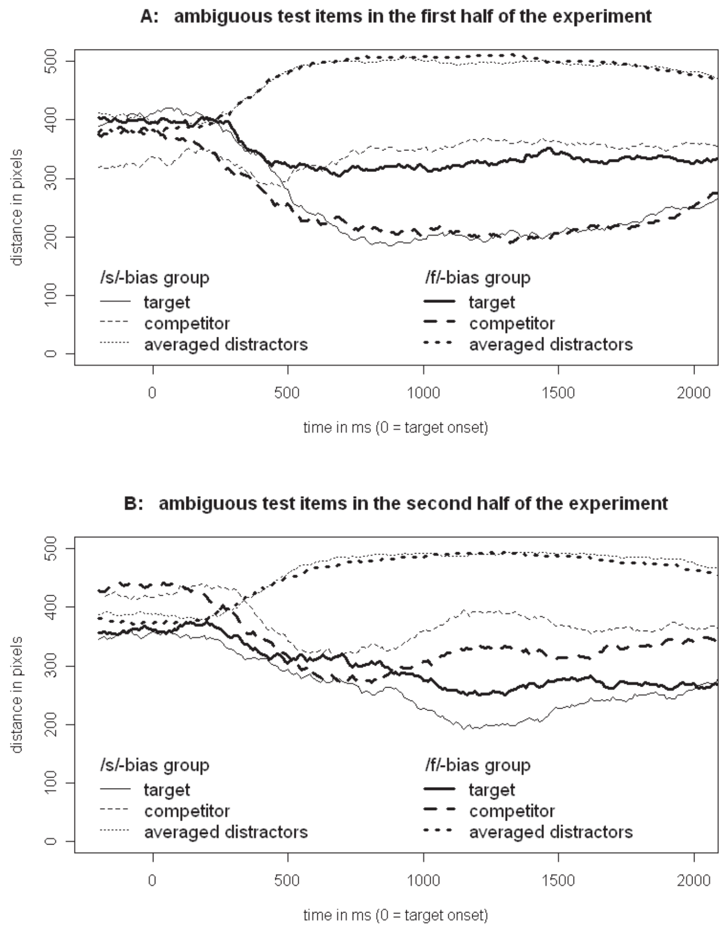
A linear mixed-effects model for the accuracy of the click responses showed an interaction of Group and Block. Separate analysis showed that the /s/-bias group performed above chance overall (68% correct responses) but did not improve over the course of the experiment. The /f/-bias group showed a significant improvement ($p < 0.05$) from 20% to 70% correct responses from the first to the last miniblock.

Eye-tracking Data

Figures 2A and 2B show the eye-tracking results for the test trials of the first and second half of the experiment. Both groups heard members of /s/-/f/-minimal pairs pronounced with an ambiguous fricative (e.g., *doo[$^{s}/_{f}$]*). In the first half of the experiment (see Figure 2A), the /s/-bias group (represented by the thin lines) started to prefer their target words (e.g., *doos*) approximately 500 ms after target onset. At this point, the target and the competitor lines diverge. The /f/-bias group, however, behaved in exactly the same way. When listeners in the /f/-bias group heard *doo[$^{s}/_{f}$]*, they also tended to look to *doos* (their competitor) and not to the target *doof*. In the second half of the experiment (see Figure 2B), a major change in the behavior of the /f/-bias group (but not the /s/-bias group) can be observed. After first considering the competitor as a possible candidate, participants in the /f/-bias group prefer to look to their target (e.g., *doof*) from 900 ms onwards.

This is borne out by statistical analyses using linear mixed-effects models with Target Preference as dependent and Group and Block as independent variables. We used sliding 200 ms time windows from 400 ms to 1600 ms after target onset. There were clear group effects in all time windows ($p_{max} = 0.0056$), indicating the better performance of the /s/-bias group, and effects of Block ($p < 0.05$) from 1000 ms till 1500 ms, indicating an overall learning effect. The learning effect was numerically larger for the /f/-bias group, confirming a similar trend as was found in the click responses. However, an interaction of Group and Block (indicating significantly more learning in the /f/-bias group) was only marginal in the late time windows.

Figure 2. Distance of participants' looks to the target, competitor, and averaged distractors in pixels during the first half (Panel A) and second half (Panel B) of the experiment. Note that looks to an object lead to diminishing distances, so that smaller distances mean more looks.



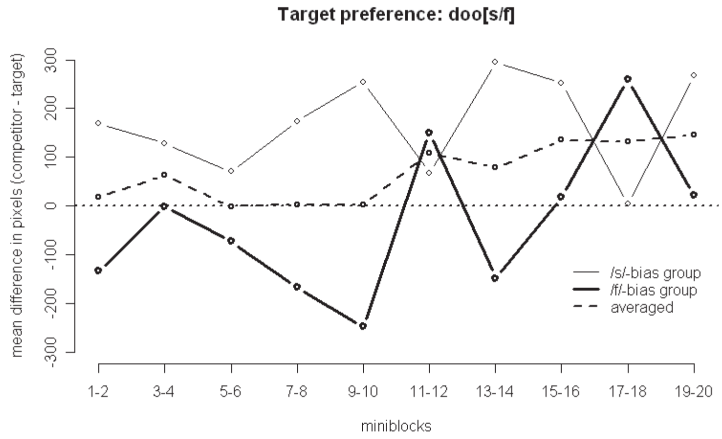
Learning function

To calculate the learning function, we subtracted for each miniblock the target fixation distance from the competitor fixation distance in the time window from 1000 ms to 1500 ms after target onset and took the average value as representative for that miniblock. If participants have a preference for the target, this difference (competitor - target) will be positive.

Figure 3 shows the learning function averaged over two miniblocks for both groups separately as well as averaged over the two groups. Given that the target preference of the /s/-bias group was above chance throughout the experiment, participants in this group did well right from the beginning and improved only slightly. Apparently, there was no reason for them to adapt to the ambiguous sound as they already perceived it as /s/ in early miniblocks. The /f/-bias group started off quite badly, but improved during the experiment. That the two functions mirror each other so strongly is evidence for stimulus-specific biases in our materials. The auditory test stimuli used for each miniblock were identical for all participants, so a strong difference between the two groups indicates that the two $[^{s}/_{f}]$-bearing test items contributing to a given data point in Figure 3 either sound quite /s/-like or quite /f/-like. Overall, the results show that most items sounded more like /s/ with the exception of blocks 11/12 and 17/18. We thus observed an overall /s/-bias in our stimuli despite the fact that we conducted two pretests that should prevent this. This probably means that there are large individual differences in how listeners perceive ambiguous sounds. Note that, as training and test trials were interwoven in this design, it was not possibly to exclude participants who did not accept the ambiguous sound as an instance of either /s/ or /f/ (depending on the group they were in) as Eisner and McQueen (2005), Norris et al. (2003) and Sjerps and McQueen (2010) have done.

The overall perceptual learning effect (driven mostly by the learning of the /f/-bias group) becomes evident after miniblock 10, as can be seen from the dashed line. The function rises quite suddenly at this point from the 0-line to values around 100 pixels and then does not rise much more. This suggests that perceptual learning occurs in a rather step-wise fashion.

Figure 3. Learning function averaged over two miniblocks for the /s/-bias group, the /f/-bias group, and averaged over both groups. Positive values on the y-axis (competitor - target) indicate more looks towards the intended target.



Discussion

The aim of this study was to replicate perceptual learning effects with an eye-tracking paradigm and hence to investigate how learning emerges over time. As in Eisner and McQueen (2005), Norris et al (2003), and Sjerps and McQueen (2010), we were able to observe a learning effect after 20 critical items, as participants had heard 10 training items but also 10 [$^{s}/_{f}$]-bearing members of temporary minimal pairs (e.g., *gister - giftig*) after miniblock 10. However, in order to replicate the findings of (Kraljic & Samuel, 2007), we should have found a learning effect from miniblock 5 onwards. A possible explanation for this difference may be that the temporary minimal pairs do not constitute training items, because there is insufficient lexical bias to guide perceptual learning at the moment the ambiguous fricative is heard. That is, the fragment *gi[$^{s}/_{f}$]* could contain an /s/ (continuing as *gister*) or an /f/ (continuing as *giftig*).

We can provisionally answer the two questions we posed. First, we found that learning apparently needs at least 10 clear training items to arise. Second, it seems that learning occurs in a step-wise fashion, as the learning effect that arose after 10 miniblocks did not get stronger with additional training.

A caveat, however, is that we observed quite a small learning effect. At least in cross-modal priming, learning appears to be "complete" after 20 training items (Sjerps & McQueen, 2010). In our data, however, the /f/-bias group still has the tendency to look at the /s/-words when hearing an ambiguous fricative, suggesting incomplete learning. A possible explanation for this discrepancy is that participants may have been aware of the purpose of the experiment. In a quarter of all trials, they saw a display like that in the right panel of Figure 1 and had to click on a member of a minimal pair. Thus, they may have paid more attention to the critical sounds and made conscious metalinguistic judgments about them. Put differently, the early test trials may have interfered with the learning process.

An interesting conclusion nevertheless arises from the fact that learning effects only occur late within a trial (> 900 ms after target onset). This suggests that perceptual learning may not influence first-pass perceptual processing, but only a reevaluation process when the input is ambiguous.

Perceptual adaptation to segmental and syllabic reductions in continuous spoken Dutch

Chapter 3

Poellmann, K., Bosker, H. R., McQueen, J. M., & Mitterer, H. (under revision). Perceptual adaptation to segmental and syllabic reductions in continuous spoken Dutch. *Journal of Phonetics*.

Abstract

This study investigates if and how listeners adapt to reductions in casual continuous speech. In a perceptual-learning variant of the visual-world paradigm, two groups of Dutch participants were exposed to either segmental (/b/ → [ʋ]) or syllabic (*ver*- → [fː]) reductions in spoken Dutch sentences. In the test phase, both groups heard both kinds of reductions, but now applied to different words. The segmental reduction exposure group was better than the syllabic reduction exposure group in recognizing new reduced /b/-words. Vice versa, the syllabic reduction group showed a greater target preference for new reduced *ver*-words. Learning about reductions was thus applied to previously unheard words. This lexical generalization suggests that mechanisms compensating for segmental and syllabic reductions take place at a prelexical level. The results suggest that lexical access involves an abstractionist mode of processing. Existing abstractionist models need to be revised, however, as they do not include representations of sequences of segments (corresponding e.g. to *ver*-) at the prelexical level.

Introduction

Phonological reductions, that is, “the articulatory weakening or complete deletion of segments and syllables” (Ernestus, 2009, p. 1875), are very common in casual speech (e.g., Johnson, 2004; Patterson, LoCasto, & Connine, 2003). A native speaker of Dutch might for instance articulate the /b/ in *baron* 'baron' as a labio-dental approximant ([ʋ]) or shorten the prefix *ver*- in *verlangen* 'desire' to [fː]. In the case of *ʋaron,* the word-initial voiced stop is weakened as it becomes an approximant and therefore less consonant-like (see definition of "lenition" in Trask, 2000). Despite these deviations from canonical pronunciations, speakers and listeners do not seem to be hampered in their communication. It is thus an intriguing question how listeners overcome such distortions in order to be able to recognize the originally intended words. The present study tests whether listeners can adapt to different types of reduction. Having heard a sufficient number of either segmental (/b/ → [ʋ]) or syllabic (*ver*- → [fː]) reductions, participants might be able to transfer their acquired knowledge about the speaker's pronunciation habits to previously unheard words.

Listeners have different ways of dealing with reduced speech at their disposal. For example, they take word frequency into account. In casual speech, a whole variety of reduced forms can be found ranging from assimilations (e.g., [tœy**m**bɑŋk] for Dutch *tuinbank* ‘garden bench’; see e.g., Mitterer & Blomert, 2003) to extreme reductions (e.g., [ɛik] instead of [ɛixələk] for Dutch *eigenlijk* ‘actually’; see e.g., Brouwer, 2010; Ernestus, 2000). Segmental and syllabic reductions lie in between these two endpoints of the reduction continuum. This continuum is determined, in part, by a word's frequency of occurrence. Usually, the more frequent a lexical item is, the more likely it is to observe strongly reduced forms (e.g., Ernestus, 2000; Jurafsky, Bell, Gregory, & Raymond, 2001). So, while any word in a sentence can undergo assimilation, only high-frequency words are reduced more severely (deletion of segments or even syllables).

It is thus not surprising that word frequency is one of the major cues that listeners use to compensate for extreme reductions like [ɛik] (e.g., Ernestus, 2000). But the sentence context also plays an important role in recognizing these forms (e.g., Brouwer, Mitterer, & Huettig, 2013; Ernestus, Baayen, & Schreuder, 2002). Assimilated forms like [tœy**m**bɑŋk], in contrast, can be compensated for by means of fine phonetic detail in the speech signal (Gow, 2002, 2003). For instance, an assimilated labial segment like the [m] in *gardem bench* is acoustically different from the intended [m] in *the same bench*.

Fine phonetic detail can hint at the presence of an apparently deleted segment, like a schwa: Listeners are able, for example, to differentiate a reduced form [spɔːt] of 'support' from the unreduced form [spɔːt] 'sport' (Manuel, 1992). However, Connine, Ranbom, and Patterson (2008) also emphasize the role of stored pronunciation variants in recognizing schwa-deleted words. In a syllable judgment and a lexical decision task, listeners reacted faster to more frequent surface forms, regardless of whether they were schwa-bearing or not. Compensation mechanisms for the deletion of word-final /t/ in Dutch (Mitterer & Ernestus, 2006) involve not only the exploitation of fine phonetic detail and lexical information but also the utilization of probabilistic cues from the surrounding context (Mitterer & McQueen, 2009b). In Dutch, when the word-final /t/ is preceded by an /s/ and followed by a bilabial sound (/b/, /m/), the /t/ is very likely to be reduced. Participants were found to be sensitive to these probabilistic facts.

Finally, the fact that assimilated forms can be recognized when listening to an unfamiliar language (Mitterer, Csépe, & Blomert, 2006; Mitterer, Csépe, Honbolygo, & Blomert, 2006) suggests that language universal auditory processes are involved in compensation, at least for some forms of assimilation (see also Steriade, 2001).

There are thus different kinds of information used when compensating for (different kinds of) reductions. Frequency information, contextual cues, fine-phonetic detail, and lexical and probabilistic knowledge all have roles to play in the compensation process, and that process is based on low-level (auditory), prelexical and lexical mechanisms.

Research on how listeners are able to compensate for reduced speech has concentrated so far on these language- and speech-dependent properties. In contrast, the current study focuses on speaker-dependent properties of reduced speech. This may be an important piece of the puzzle, since reductions depend on the speaker in several ways: First, reductions are optional. There are of course phonological contexts in which they are more likely to occur, but nothing restrains the speaker from articulating words in their full form (Mitterer & Ernestus, 2006). Similarly, reductions are not a direct by-product of a fast speaking rate. There are speakers who speak very fast but clearly (Shockey, 2003; Van Son & Pols, 1990, 1992), while others even produce reductions when reading aloud (Warner & Tucker, 2011). Second, when speakers reduce words, they can do so in many different ways. For instance, the Dutch suffix *-lijk* can be realized at a continuum ranging from the citation form [lək] to highly reduced [ə] or [k] (Pluymaekers, Ernestus, & Baayen, 2005). As speakers can choose from a great variety of possible reductions of a given sequence, they vary in the choices they make (Ernestus, et al., 2002). Moreover, Keune, Ernestus, van Hout, and Baayen (2005), investigating how common reductions are among different talker groups (male and female speakers of Dutch and Flemish), found that Dutch men reduced most often, while Flemish highly educated women were least likely to do so.

The question then arises whether listeners tune in to the reduction styles of given speakers, assuming that these speakers reduce words consistently, in order to compensate for their sloppy speech. In other words, is there adaptation to specific reduction styles? We asked that question here. In particular, we tested whether adaptation to a specific reduction facilitates the recognition of new reduced words spoken by the same talker.

There is ample evidence that listeners tune in to characteristics of particular speakers. For instance, Sjerps, Mitterer, & McQueen (2011) found that the same vowel ([ɪ] or [ɛ]) was perceived differently depending on the F1 of the context it appeared in. So listeners do not interpret an incoming sound solely on the basis of its acoustic properties, but take speaker-specific properties (like the available F1 range) into account. Another example of how listeners tune in to a particular talker is provided by Reinisch, Jesse, and McQueen (2011). They showed that the perception of a juncture phoneme with a constant, but ambiguous duration (e.g., [s] in Dutch 'eens (s)peer', "once (s)pear") changed depending on the rate at which the preceding context was spoken. Again, listeners do not just evaluate the raw incoming signal, but they compute the incoming information (segment duration in this case) relative to talker-specific information (like speaking rate) available from the preceding context.

These normalization processes are assumed to occur at an early auditory level of processing (e.g., Reinisch & Sjerps, 2013; Sjerps, et al., 2011). In line with this assumption, normalization for speaking rate is speaker-independent, so that speaker identity of the precursor has hardly any effects on normalization (Newman & Sawusch, 2009). Another normalization process, however, is more tightly coupled to speaker identity. Norris, McQueen, and Cutler (2003) tested whether listeners could adapt to speaker-specific idiosyncrasies, in their case a speaker who produced either the fricative /f/ or the fricative /s/ as an ambiguous sound ([$^{s}/_{f}$]). In these perceptual learning studies (for a review, see Samuel & Kraljic, 2009), one group of participants was trained to perceive an ambiguous sound [$^{s}/_{f}$] as /s/, another group was trained to perceive the same sound as /f/ based on lexical bias. This kind of exposure leads to a recalibration of the /s/-/f/ contrast which generalizes across words (McQueen, Cutler, & Norris, 2006). Moreover, this perceptual learning effect proved to be talker-specific and stable over time (Eisner & McQueen, 2005, 2006). Investigating whether there is "a return to normal", Kraljic and Samuel (2005) found that only canonical pronunciation of *both* phonemes (/s/ and /ʃ/ in this case) appeared to be able to reset the phonemic categories to pre-learning parameters. Importantly, these unambiguous instances had to be uttered by the same speaker listeners had been trained on. Learning about segmental idiosyncrasies involving fricatives (i.e., sounds that convey talker-specific information) was again found to be talker-specific, while learning about plosives (which contain hardly any talker-specific information) generalized across speakers (Kraljic & Samuel, 2007). In all these studies, listeners learned to interpret an artificially constructed ambiguous sound as an (odd) pronunciation variant of a given phoneme (depending on the use the speaker made of this sound).

Maye, Aslin, and Tanenhaus (2008) showed that listeners can also adapt to vowel differences that depend on sociolinguistic factors. In their study, they tested adaptation to a hypothetical regional dialect of American English. Participants listened to a story twice (in two separate sessions), first to a canonical version and the second time to a manipulated version in which the front vowels were lowered in F1-F2 vowel space. In the following lexical decision task, which was identical for the two sessions, listeners categorized more accented words as existing English words after exposure to the accented story than after exposure to the canonical version of the story. This adaptation effect generalized to words which had not occurred in the story.

Dahan, Drucker, and Scarborough (2008) investigated adaptation to a natural dialect of American English. Participants were exposed to a dialect speaker who naturally raised the vowel /æ/ before /g/ but not before /k/. They could identify *back*-like words (which are unaffected by the dialect shift) more easily after exposure to accented *bag*-like words. This facilitation was probably due to reduced competition between the two word types because the different vowels contributed to the disambiguation. Listeners can thus adapt to unusual pronunciations attributed to the dialect of a given speaker.

In the present study, we asked whether listeners can also adapt to the reduction style of a given speaker. That is, like Dahan et al. (2008), we wanted to test listeners' ability to adjust to naturally-occurring idiosyncrasies of a speaker, but unlike them we were interested in stylistic rather than in geographical variation as the source of the idiosyncrasies.

In the present experiments, we used a variant of the visual-world eye-tracking paradigm (Mitterer & McQueen, 2009b) that was modified to permit study of perceptual learning. Different groups of participants were exposed to one of two different forms of reduction, and then tested on both forms. One group of participants was exposed to segmental reductions; the other group to syllabic reductions. That is, one group (the segmental reduction group) heard /b/-initial words where the /b/ of an unstressed first syllable was replaced by a sound similar to the Dutch labio-dental approximant [ʋ] (e.g., [ʋikini] instead of [bikini] 'bikini'). The other group (the syllabic reduction group) heard *ver*-words where the prefix *ver*- was replaced by a long [f] (e.g., [fːledə] instead of [vərledə] *verleden* 'past'; note that most Dutch speakers already devoice initial /v/ to [f], so that the latter form becomes [fərledə]). In the test phase, both groups heard both kinds of reductions. We reasoned that, if listeners are able to tune in to respective reduction styles, the segmental reduction group should be better at recognizing new reduced /b/-words than their control group (the syllabic reduction group), and the opposite should be true for new reduced *ver*-words (the syllabic reduction group should perform better on them than the segmental reduction group).

In the test phase, all reduced items were words that participants had not encountered in the exposure phase. Finding an adaptation effect would thus be evidence of generalization of learning to new words. Such a finding would contribute to the ongoing debate on theories of speech processing. In some models (e.g., TRACE, McClelland & Elman, 1986; Shortlist B, Norris & McQueen, 2008), lexical representations are abstract: The acoustic input is mapped onto abstract phonological units and then onto corresponding entries in the mental lexicon. Prelexical processes have to compensate for continuous speech phenomena like reduction so that underlying phonological representations can be accessed. In contrast, in episodic models, the entry for a particular word in the mental lexicon consists of detailed and concrete episodes of a word's pronunciation that the listener has encountered previously (e.g., Bybee, 2001; Goldinger, 1998; Hawkins, 2003). Thus, an entry not only consists of various memories of a word's citation form, but also contains episodic traces of reduced forms showing different kinds and/or degrees of reduction. A model storing only acoustically detailed lexical episodes and using only those episodes in word recognition cannot take advantage of any sublexical regularities during the word recognition process.

It has been shown that neither purely abstractionist models (e.g., Nygaard & Pisoni, 1998) nor extreme episodic models (e.g., Cutler, Eisner, McQueen, & Norris, 2010) can explain all aspects of spoken word recognition alone. The field, therefore, has converged on the idea of hybrid models combining aspects of both theories (e.g., Cutler, et al., 2010; Ernestus, 2009; Mitterer, Chen, & Zhou, 2011). In this context, we can ask how important the episodic and abstractionist modes of lexical access are in the processing of reductions. If episodic storage of reduced forms is the only means to recognize them, there should be no generalization of learning about reductions to new reduced words. If, however, abstraction plays an important role in the recognition of reduced forms, such generalization should occur. Abstraction would enable the learning effect to generalize to previously unheard reduced words.

Our experimental set-up, however, may also provide data that are problematic for current versions of abstractionist models. This is because we test whether listeners can adapt to regular syllabic reductions, such as the Dutch prefix *ver*- being shortened to a long [f] (e.g., [fːledə] for *verleden*). In current abstractionist models, the acoustic input is mapped at the prelexical level onto abstract phonological representations. These are taken to be features (Gaskell & Marslen-Wilson, 1997), phonemes (Norris & McQueen, 2008) or features and phonemes (McClelland & Elman, 1986). None of these models, however, stipulates prelexical representations that are larger than a segment and that would be able to encode at a prelexical level that [f] is a possible pronunciation of the three-segment string /vər/. Such representations, however, would be necessary to explain generalization from one word to another for a reduction such as /vər/ → [f].

Before testing whether listeners can adapt to reductions, we needed to know whether Dutch speakers actually produce the kind of /b/- and *ver*-reductions we wanted to investigate. Pluymaekers et al. (2005) examined reduced affixes in spoken Dutch, amongst which the prefix *ver*-. They first established how many different word types starting with *ver-* occurred in the Corpus of Spoken Dutch (Oostdijk, 2000). Here, the notion "word type" not only includes words belonging to different lemmas but also different word forms of one lemma (e.g., the present and past tense of a given verb). Only one randomly chosen token per word type was considered for analysis and thus not all instances of *ver*-words that occurred in the corpus. The prefix *ver-* was reduced to [f] in 23 out of 140 word types (i.e., in 16.43%). There were no such data available on word-initial /b/-reductions resulting in a labio-dental approximant. So, following Pluymaekers et al. (2005), we conducted a similar study for the /b/-reductions making use of the same corpus. Out of 111 word types of /b/-initial words, where the /b/ was followed by a full vowel and where the first syllable was unstressed, 15 were pronounced with a labio-dental approximant [ʋ]. This was the most frequent non-canonical variant (13.51%). Its frequency is similar to the percentage of *ver*-to-[f]-reductions. Word-initial /b/s in unstressed position may also be reduced to approximants in other languages. Warner and Tucker (2011), investigating phonetic variability of stops in spontaneous and careful speech, noted for American English that expected voiced stops (/b, ɡ/) may be realized as approximants or vowel-like sounds. All stops examined in their study occurred between vowels and/or sonorants. Consonants were more reduced between two unstressed syllables (also across word boundaries) than after a stressed syllable. In the current study, the experimental sentences were chosen to fulfill all of these conditions.

Experiment 1

The aim of Experiment 1 was to apply a learning paradigm consisting of exposure and test phases to investigate whether listeners are able to adapt to talker-specific reduction styles. In the exposure phase, one group of participants (the segmental reduction group) was exposed to /b/-reductions, while another group (the syllabic reduction group) was exposed to *ver*-reductions. In the test phase, both groups were tested on new reduced words of both reduction types. If participants in the segmental reduction group can tune in to /b/-reductions, they should perform better than the syllabic reduction group on new reduced /b/-words. The same holds for participants in the syllabic reduction group: If they can adapt to *ver*-reductions, they should outperform participants in the segmental reduction group on new reduced *ver*-words.

Method

Participants. Fifty-six participants from the Max Planck Institute subject pool, all native speakers of Dutch, were paid to take part. None reported any hearing disorders and all had normal or corrected-to-normal vision.

Design. In the exposure phase, one group of participants (the segmental reduction group) was exposed to four critical types of items (see Table 1): They heard reduced /b/-words (e.g., /bikini/ *bikini* produced as [$^{b}/_{ʋ}$ikini]), canonical /ʋ/-words (e.g., /ʋustənɛɪ/ *woestenij* produced as [ʋustənɛɪ]), canonical *ver*-words (e.g., /vərledə/ *verleden* produced as [fərledə]) and quasi-canonical /f/-words (e.g., /flɔsə/ *flossen* produced as [fːlɔsə]). The term *quasi-canonical* is used for pronunciations that deviate slightly from the canonical form, but are not reduced variants. The other group (the syllabic reduction group) was exposed to the reversed design: They heard canonical /b/-words, quasi-canonical /ʋ/-words (e.g., /ʋustənɛɪ/ *woestenij* produced as [$^{b}/_{ʋ}$ustənɛɪ]), reduced *ver*-words (e.g., /vərledə/ *verleden* produced as [fːledə]) and canonical /f/-words. This procedure was chosen to enhance the contrast between the two groups. Listeners in the segmental reduction group were led to interpret the ambiguous sounds [$^{b}/_{ʋ}$] as [b] (hearing them at the beginnings of /b/-words) and were encouraged to disregard the cues indicating a syllabic reduction (hearing the long [f]s in /f/-words). Conversely, the syllabic reduction group learned that a long [f] is a variant of *ver*- that the speaker produces, and were encouraged to ignore the cues indicating a /b/-reduction (hearing the [$^{b}/_{ʋ}$]s at the beginnings of /ʋ/-words).

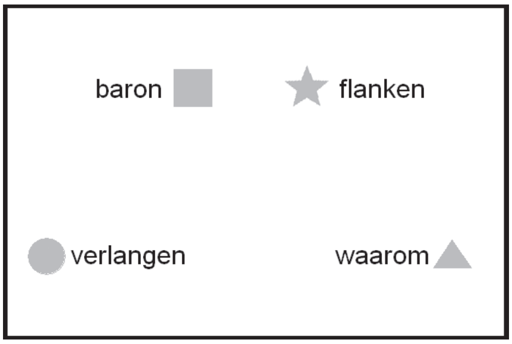
In the test phase, both groups heard only reduced forms of both types, in new /b/-, /ʋ/-, *ver*- and /f/-words (see Table 1). If participants did not adapt to the reduction type they were exposed to, there should not be a group difference in the test phase. If, however, participants did adapt to the reduction type they were exposed to, they should outperform the other group on new items of the same reduction type. Thus, each group of participants served as an experimental group in one condition and as a control group in the other.

Table 1. Experimental design and types of stimuli.

	/b/-reduction group (segmental)		*ver*-reduction group (syllabic)	
Exposure phase	reduced /b/-words	[$^{b}/_{ʋ}$]ikini	canonical /b/-words	[b]ikini
	canonical /ʋ/-words	[ʋ]oestenij	quasi-canonical /ʋ/-words	[$^{b}/_{ʋ}$]oestenij
	canonical *ver*-words	[fər]leden	reduced *ver*-words	[fː]leden
	quasi-canonical /f/-words	[fː]lossen	canonical /f/-words	[f]lossen
Test phase	all items reduced	[$^{b}/_{ʋ}$]aron	all items reduced	[$^{b}/_{ʋ}$]aron
		[$^{b}/_{ʋ}$]aarom		[$^{b}/_{ʋ}$]aarom
		[fː]langen		[fː]langen
		[fː]lanken		[fː]lanken

Materials. The critical words were embedded in a spoken carrier sentence like "Klik op het woordje XXX daar links van het rooie vierkantje" ('Click on the word XXX there to the left of the red rectangle'). In this sentence, the word preceding the target word (i.e., the word which participants should click on) ended in a sonorant voiced segment (the schwa of *woordje* [ʋoːrtʃə]). This is the phonological context for the occurrence of /b/-reductions (Warner & Tucker, 2011). Note that *ver*-reductions occur in a variety of contexts. Participants hearing a sentence with the reduced word *[$^{b}/_{ʋ}$]aron* 'baron' would see a display like the one in Figure 1. In the exposure and test trials, there were always a /b/-word, a /ʋ/-word, a *ver*-word and an /f/-word on the screen.

Figure 1. Example display of a test trial.



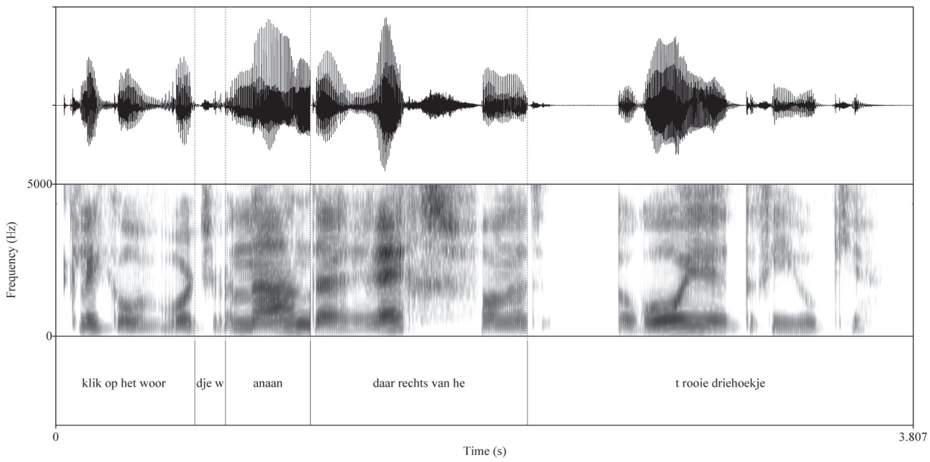
For the test phase of the segmental reduction condition, 24 /b/-initial words were selected which did not carry initial stress. Their /ʋ/-initial competitors were matched on following vowel and stress pattern (e.g., boerenKOOL 'kale' and woesteNIJ 'wasteland' both showing a weak-weak-strong pattern). For the test phase of the syllabic reduction condition, 24 words beginning with *ver-* and followed by a liquid were selected as targets. Their competitors were /f/- or /v/-initial words followed by an identical combination of liquid and vowel, so that a target-competitor pair would overlap in at least three segments in their reduced forms (e.g., the first syllable(s) of *verliezer* 'loser' and *vlieger* 'kite' would both be pronounced as [fli]). The target-competitor pairs were also matched on number of syllables in their reduced forms (i.e., *verliezer* is counted as having two syllables when reduced and thus matches *vlieger*). However, due to the onset requirements of the stimuli as well as their stress pattern constraints, it was not possible to match the stimuli on frequency of occurrence (see Appendix A, Table A1). Furthermore, an additional 24 /ʋ/-initial and 24 /f/- or /v/-initial target words as well as their /b/- and *ver*-initial competitors were selected in order to balance the responses to /b/- and /ʋ/-words as well as the responses to *ver-* and /f/-words.

For the exposure phase, four new sets of 24 /b/-, /ʋ/-, *ver-*, and /f/- or /v/-initial words and competitors were selected. The target exposure words did not overlap with their competitors to the same extent as the test words did.

Stimulus construction. Digital recordings of the stimuli were made by a male native speaker of Dutch (the second author) in a sound-proof booth, sampling at 44.1 kHz. The speaker naturally devoices instances of [v] to [f], so that a *ver*-to-[f]-reduction was plausible. As a consequence, it was possible to use /v/-initial words (which would be devoiced by this speaker) and /f/-initial words as competitors for the *ver*-words. In order to make the occurrence of reductions more plausible, we tried to elicit a casual speaking style by asking the speaker to produce sentences according to the following template: "Klik op het woordje ... daar links/rechts van het gele/groene/rooie/zwarte rondje/sterretje/driehoekje/vierkantje" ('Click on the word ... there to the left/right of the yellow/green/red/black circle/star/triangle/rectangle'). The sentences were prompted by a visual display on a computer screen with a printed target word and the corresponding shape positioned to the left or to the right of the word. This recording method produced some hesitations, which were retained as they made the recordings sound more casual. Also the colloquial *daar* 'there' was included in the sentence for this purpose. The sentence accent was always placed on *rechts* or *links*. As for the target words, the speaker produced each /b/- and *ver*-word both canonically and in reduced form, that is, he consciously replaced the initial [b]s by short [ʋ]s and the initial *ver*-s by long [f]s.

The experimental sentences were created by cross-splicing. Splicing points were chosen at major positive-going zero-crossings before the stops in *woordje*, *daar* and *het* (see Figure 2). The first fragment *Klik op het woor*- was the same for all sentences. Likewise, only one version of the sentence parts *daar links van het*, *daar rechts van het*, and of each color-shape combination was used.

Figure 2. Splicing points in an experimental sentence containing a segmental reduction in Experiment 1.



To create the reduced /b/-words, ambiguous sounds [$^{b}/_{ʋ}$] midway between a /b/ and a /ʋ/ were selected by means of a pretest (see below). These ambiguous sounds and their respective *dje's* were then spliced between the first sentence fragment (*Klik op het woor*-) and the rest of the /b/-word respectively (see Figure 2). For the reduced *ver*-words, the recorded versions containing a long [fː] were used. That is, the reduced *ver*-words were not cross-spliced. The canonical versions of the /ʋ/-initial and the /f/-initial words were cross-spliced to create what we called the quasi-canonical forms. For the

/ʋ/-initial words, the same procedure as for the reduced /b/-words was used. For the /f/-initial words, one of the *dje* [fː]-sequences (i.e., the consciously reduced *ver*- and its original *dje*) was spliced between the first sentence fragment (*Klik op het woor*-) and the rest of the original /f/-word. As the splicing point in the /f/-initial words was located at the end of a fricative and the spliced-in segment ended in a fricative as well, the transition did not give the splicing away. The remainder of the /f/-word fitted well with a somewhat longer [f] than the one of the original recording. It thus did not matter that /f/-initial words were manipulated while reduced *ver*-initial words were not.

In order to prevent coarticulation features in the cross-spliced materials from interfering, attention was paid to the fact that the source words of the reduced segments (the various [$^{b}/_{ʋ}$]s and [fː]s) overlapped in at least one following segment with the target words they were spliced onto (e.g., the ambiguous [$^{b}/_{ʋ}$] that stemmed from the consciously reduced *baron* was also used for *waarom* and *barak*). All /b/- and /ʋ/-initial target words of the test phase therefore started with an ambiguous [$^{b}/_{ʋ}$]. Similarly, all *ver*- and /f/-initial test targets started with a long [fː]. In the exposure phase however, the segmental reduction group heard manipulated /b/-words (containing [$^{b}/_{ʋ}$]s) and /f/-words (containing [fː]s) only, while the syllabic reduction group heard manipulated *ver*-words (containing [fː]s) and /ʋ/-words (containing [$^{b}/_{ʋ}$]s) only.

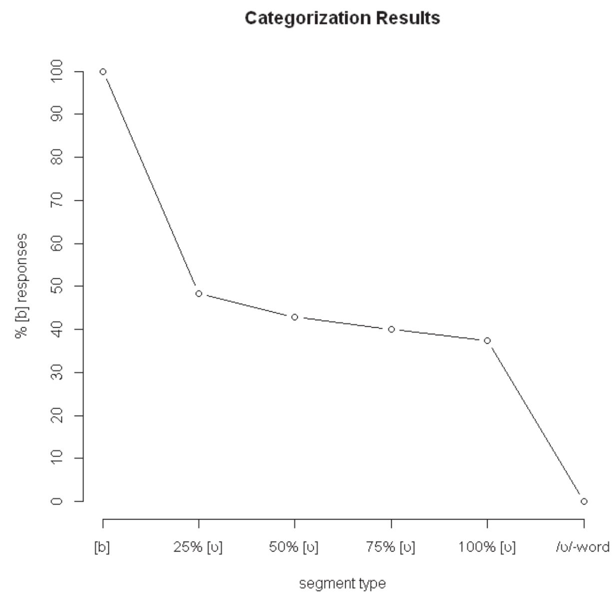
For the /b/-reductions, an ambiguous sound midway between /b/ and /ʋ/ was chosen in order to stay as close as possible to the classical perceptual learning paradigm by Norris, et al. (2003). The reduction of the prefix *ver*-, however, which resulted in a single fricative (/f/), was less ambiguous. Therefore, the fricative was lengthened in order to provide listeners with fine phonetic detail which they could use for learning.

Pretest. In order to find ambiguous sounds [$^{b}/_{ʋ}$] between /b/ and /ʋ/ that would serve as reduced [b]s, a pretest was run. Five native Dutch employees of the Max Planck Institute for Psycholinguistics, Nijmegen, participated.

A /b/-/ʋ/-continuum was created by excising each first segment of the 48 consciously reduced /b/-initial words (i.e., the /b/-initial words that were recorded with a [ʋ] instead of a [b]) and shortening its original duration to 25%, 50%, and 75% with PSOLA (using Praat 5.1.26, Boersma & Weenink, 2010). Five versions of every /b/-initial word (e.g., *banaan* 'banana'; one version pronounced canonically with [b], one pronounced with [ʋ]—*wanaan*—plus the three shortened versions of [ʋ]) as well as the 48 canonically pronounced /ʋ/-initial words (e.g., *waanzinnig* 'insane') were presented to the participants in a random order which was different for each participant. Participants heard sentences like "Klik op het woordje banaan" ('Click on the word banana') and had to decide whether the last word of the sentence was pronounced with a [b] or with a [ʋ], regardless of the spelling of the word. Participants were told explicitly that /b/-initial words could be mispronounced with an initial /ʋ/. They were not put under time pressure and were encouraged to make their decision carefully. After the sentence was played, they were asked to click with the mouse on a field labeled "b" or "w" (the orthographic form of /ʋ/ in Dutch) depending on what they had heard.

Figure 3 shows the results of the pretest. Quite unsurprisingly, the canonical /b/- and /ʋ/-words got 100% and 0% of [b] responses respectively. However, the different replacements of the [b] in the /b/-words were all judged quite similarly with the votes lying within a range from 49% to 37% of [b] responses. Although the value of the 25% [ʋ] stimuli was closest to the 50% criterion (i.e., the perfect value of ambiguity), we decided to use the 50% [ʋ] stimuli for the main experiment, as their values were comparable to those of the 25% [ʋ] stimuli but the items were less strongly manipulated and hence sounded more natural.

Figure 3. Pretest: Percentages of [b] responses for /b/-initial words (e.g., *banaan* 'banana') whose first segment was a natural [b], 25%, 50%, 75% and 100% of a [ʋ] replacing the original [b] (e.g., *wanaan*) and canonical /ʋ/-initial words (e.g., *waanzinnig* 'insane').



Procedure. Participants were seated in a sound-attenuated booth at a comfortable viewing distance from the computer screen. Eye movements were monitored using an SR Research EyeLinkII, sampling at 500 Hz. The auditory stimuli were presented to the participants over headphones. Prior to the experiment, participants received written instructions in which they were shown an example sentence and an example display. The instructions read that they should click on the word that occurred in the sentence. Participants were told to pay attention to the whole sentence, as it could be that a word appeared twice on the screen (filler items). They were told that, in these cases, the disambiguating information about the color, form or position of the geometrical shape relative to the target word would be provided in the last part of the sentence.

During the experiment, participants heard sentences like "Klik op het woordje [fː]langen daar rechts van het rooie rondje" ('Click on the word desire there to the right of the red circle') and were asked to follow these instructions. Figure 1 shows an example of a display; the example is taken from a trial in the test phase. In the exposure and test trials, there were always a *ver*-word, an /f/-word, a /b/-word and a /ʋ/-word on the screen which were associated with one of four different shapes in the same color (see Figure 1). Each center of the printed words coincided with the center of one quadrant of the screen, independently from word length. The corresponding shapes were either placed to the left or to the right of the printed words. The shapes did not have an absolute position on the screen, rather the distance to the printed words was kept constant. Targets and competitors were positioned on the same side of their corresponding geometrical shape.

The main difference between exposure and test displays lay in the degree of similarity between target and competitor words. In the exposure phase, they did not share any overlapping segments following the reduced onset segment(s) (e.g., *bikini* 'bikini' and *wagon* 'wagon' or *verleden* 'past' and *vogelkooi* 'birdcage') while they overlapped at least until the first vowel following the reduced onset segment(s) in the test phase (e.g., *baron* 'baron' and *waarom* 'why' or *verliezer* 'loser' and *vlieger* 'kite').

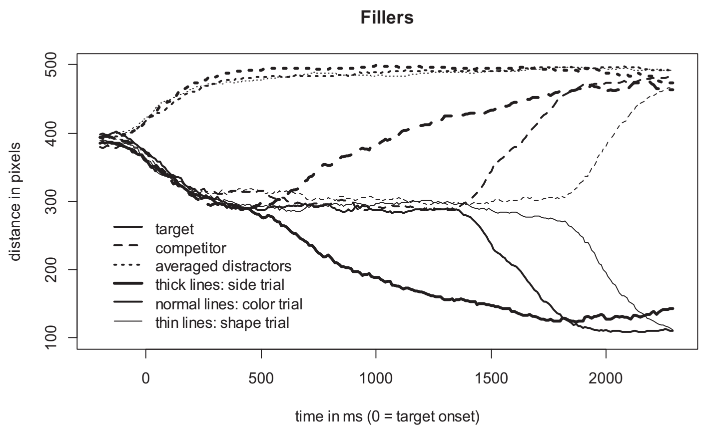
Furthermore, there were 48 filler trials interspersed among the 96 exposure and 96 test trials. Fillers consisted of two different words that appeared twice on the screen. The identical words differed either in their position (left or right) relative to the same shape in the same color (i.e., it was a side trial), or they were placed on the same side of the same shape but the color of the shape was different (a color trial), or they were positioned on the same side of different shapes in the same color (a shape trial). The filler trials were supposed to draw participants' attention to the information given in the second half of the experimental sentences (i.e., the information about position, color and shape of the object associated with the word they should click on) so that participants would not concentrate too much on the reduction phenomena. Mitterer and McQueen (2009b) found that the inclusion of such trials made phonological effects more consistent.

At the beginning of each trial, a fixation cross appeared in the center of the screen for 500 ms. Then the four printed words (in a 28-point Arial font) and the four associated shapes were presented. After 1500 ms, the auditory stimulus was played. As soon as participants had listened to the entire sentence and had clicked with the mouse on the screen, the following trial was initiated. Every ten trials, a drift correction was carried out. Participants had the opportunity to take a break after every 50th stimulus. The experiment started with four practice trials. The randomized order of presentation (exposure and filler trials, then test and filler trials) was different for each participant. An experimental session took approximately 35 minutes.

Results

Fillers. Filler trials were not considered for statistical analysis. Nevertheless, we plot the eye movement data of the fillers in Figure 4 to show how participants used the unfolding information over time. On the y-axis, the distance in pixels between participants' fixation points and the centers of the three word types on the screen (target, competitor and two distractors) are plotted (100 pixels correspond to 3.2 cm in Experiment 1; average word length was 165 pixels). If participants look to the target word, the distance between the fixation point and the center of the target word should become smaller. In the instruction sentence, the disambiguating information became available first in the side trials then in the color trials and finally in the shape trials. This is reflected in when participants started to prefer the target over the competitor: A preference for the target over the competitor arose first in side trials, then in color trials, and finally in shape trials. In a side trial (represented by thick lines), participants started to look to the target at around 500 ms after target onset. In a color trial (represented by lines of medium thickness), participants had a preference for the target about 1400 ms after target onset. Finally, participants started to look to the target only about 1800 ms after target onset if the information about the shape (thin lines) disambiguated the target word. Figure 4 thus shows that participants were using information as soon as it was available.

Figure 4. Distances between fixations and targets, competitors and averaged distractors in pixels for the filler trials in Experiment 1 (e.g., *Klik op het woordje konijn daar rechts van het groene driehoekje* 'Click on the word rabbit there to the right of the green triangle').



Exclusion criteria. The dependent variables were mouse clicks and eye movements. Trials in which participants did not click within a circle with a radius of 180 pixels around one of the centers of the four words on the screen were not considered further. Due to this criterion, 0.3% of all trials (36 trials) were excluded. For the eye-tracking data, we analyzed the data from the better eye of the participants (i.e., the eye that showed less error in the validation of the calibration of the eye-tracker). Only trials in which participants clicked on the target word were included in the eye movement analysis. This led to the exclusion of 81 trials (0.7%) in total. Out of these 81 trials, 76 were trials of the test phase in which participants clicked incorrectly on the competitor (1.4% of the test trials). Moreover, 196 trials (1.8%) in which participants did not look on the screen at some point in time were discarded.

Participants in both groups performed the task very accurately. The percentages of correct responses (clicks on the target word) equalled or exceeded 96% in all conditions for both exposure and test phases. Descriptive statistics of reaction times (RTs) for correct trials are displayed in Table 2 for both groups in the exposure and test phases. The standard deviation (sd) values indicate a high amount of between-subject variability. Some participants clicked as soon as possible, others often waited for the instruction sentence to finish. In general, participants in the syllabic reduction group responded faster than participants in the segmental reduction group (see also Figures 5 and B1). The large difference between minimal and maximal RTs prompted us to exclude outliers. To achieve this, a linear mixed-effects model containing only participants and items as random effects and Trial Number as fixed effect was run. Then the residuals of this (atheoretical) model were computed. Based on a visual inspection of a residual plot, 33 trials in the exposure phase (with residuals higher than 1800 ms or lower than -1300 ms) and 21 trials in the test phase (with residuals either below -1500 ms or above 2300 ms) were excluded.

Table 2. Descriptive statistics of RTs for correct trials for both groups in the exposure and test phases in Experiment 1.

	Segmental reduction group		Syllabic reduction group	
RT in ms	Exposure	Test	Exposure	Test
mean	2278	2362	2104	2207
sd	804	859	803	800
min	495	551	326	342
max	9177	8385	5471	7110

Statistical testing. We used linear mixed-effects models on both the click responses and the eye movement data. Analyses were conducted for the segmental reduction condition (/b/-words and /ʋ/-words) and the syllabic reduction condition (*ver*-words and /f/-words) separately. Participants and Items were entered in the model as random effects. Group and Target Word (/b/- vs. /ʋ/-target; *ver*- vs. /f/-target) served as fixed effects. Both fixed effects were coded as a numerical contrast (-0.5 and 0.5). In all following analyses, the segmental reduction group is coded as -0.5 and the syllabic reduction group as 0.5. Likewise, /b/-targets are always coded as -0.5 and /ʋ/-targets as 0.5 in the segmental reduction condition. In the syllabic reduction condition, the *ver*-targets are coded as -0.5 and the /f/-targets as 0.5. Task performance, as measured for example by RT, often improves over the course of an experiment. Therefore, Trial Number was added as another fixed effect with values centered around zero in the model for the click data. Thus, we tested whether RTs and target preference (as determined by the distance between competitor fixations and target fixations) were influenced by the above mentioned fixed effects. For the eye-tracking analyses, we used sliding 200 ms time windows from 200-1500 ms after target onset starting at every 100 ms. To estimate *p*-values, Markov chain Monte Carlo simulations were used.

Note that in the exposure phase, Group (exposure to segmental or syllabic reductions) and Target Word (e.g., /b/- vs. /ʋ/- word) are necessarily confounded, as only the segmental reduction group was exposed to reduced /b/-words and only the syllabic reduction group was exposed to reduced *ver*-words. Note also that an effect of reduction should manifest itself either as a main effect of Group or as an interaction of Group and Target Word. In the exposure phase, the syllabic reduction group, for example, should be slower to respond to *ver*-words than the segmental reduction group, as the former hears them in their reduced forms whereas the latter hears them in their canonical forms (see e.g., Pitt, 2009, on the RT advantage for canonical over reduced forms). The test phase will show whether we can find an adaptation effect, that is, an effect of exposure. This effect should manifest itself again as either a main effect of Group or as an interaction of Group and Target Word, but this time the effect should go in the opposite direction: The syllabic reduction group should respond faster to the *ver*-words than the segmental reduction group, as both groups now hear reduced *ver*-words and the syllabic reduction group has the advantage of being already familiar with *ver*-reductions. The reverse holds for segmental reductions.

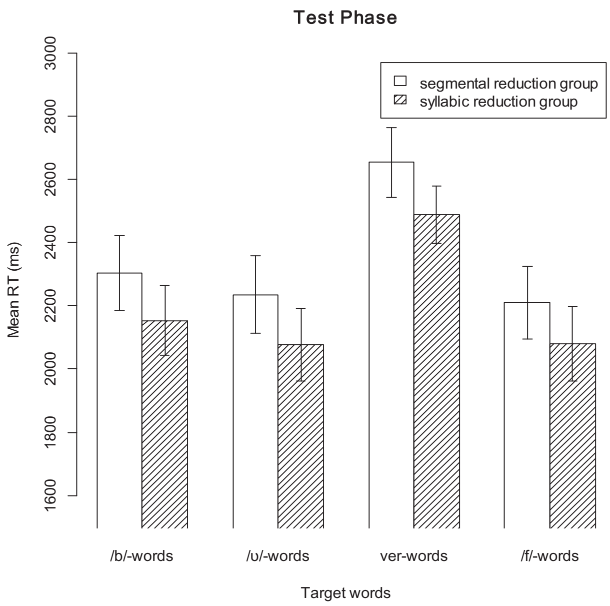
Similar results are expected for the eye movement data, with a greater preference for the target word when it was heard in its canonical form in the exposure phase and a greater preference for the target word when it was produced with the more familiar reduction type in the test phase.

Exposure Phase. Only an overview of the results will be given here. For detailed analyses, see Appendix B. Both groups reacted equally fast to /b/-words, regardless of whether they were reduced or not. The RT data showed further only a weak effect of reduction for the syllabic reduction condition. That is, participants in the syllabic reduction group were slowed down more strongly by reduced *ver*-words compared to participants in the segmental reduction group who heard canonical *ver*-words. The eye-tracking data, however, revealed significant effects of reduction for both conditions: Participants in the segmental reduction group (who heard reduced /b/-words) showed a smaller target preference for /b/-words than participants in the syllabic reduction group (who heard canonical /b/-words). The reverse was true for the syllabic reduction condition: Participants in the syllabic reduction group (who heard reduced *ver*-words) had a smaller target preference for *ver*-words than participants in the segmental reduction group (who heard canonical *ver*-words).

Test phase. The click responses for the test phase are displayed in Figure 5 in terms of mean RTs for both groups and each type of target word. In the test phase, both groups heard reduced /b/- and reduced *ver*-words, and the stimuli in the "competitor" conditions were manipulated: The [ʋ]s in /ʋ/-words were replaced by [$^{b}/_{ʋ}$]s and the [f]s in /f/-words were replaced by [fː]s. Although the syllabic reduction group was numerically faster in responding than the segmental reduction group, this difference was again not significant, as there was no main effect of Group in either of the two conditions. The main effect of Trial Number was not significant either, indicating that participants did not react faster over the course of the test phase.

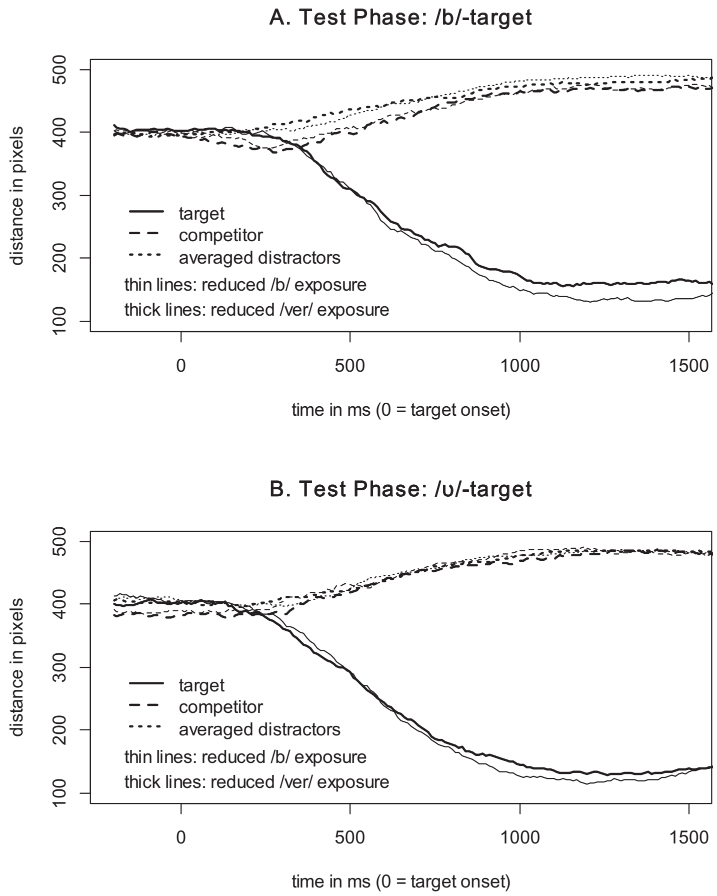
In the segmental reduction condition, the main effect of Target Word was marginally significant ($b_{\text{Target Word}}$ = -91, p_{MCMC} = 0.06). That is, both groups responded to quasi-canonical /ʋ/-words slightly faster than to reduced /b/-words. Comparing the RTs for reduced ver-words and quasi-canonical /f/-words, it can be seen that both groups took longer to respond to the reduced *ver*-words. This difference was significant (i.e., there was a main effect of Target Word: $b_{\text{Target Word}}$ = -435, p_{MCMC} < 0.001). The interaction of Target Word and Group was not significant in either condition. Thus, we did not observe an effect of exposure in the click responses of the test phase for the two conditions.

Figure 5. Mean RTs in ms and SEs of both groups for each type of target word in the test phase of Experiment 1. All /b/- and *ver*-words were reduced. Also the /ʋ/- and /f/-words were slightly manipulated: The [ʋ]s of /ʋ/-words were replaced by [$^{b}/_{ʋ}$]s and the [f]s of /f/-words were replaced by [f:]s (for more details see main text).



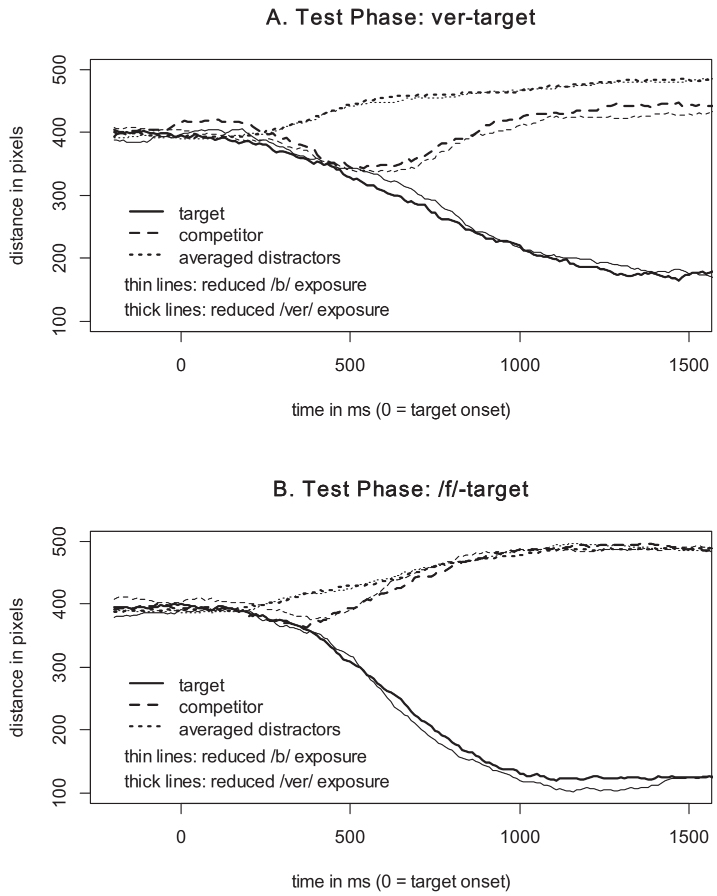
Figures 6 and 7 display the corresponding eye-tracking results of the test phase. In the segmental reduction condition (Figures 6A and B), participants in both groups behaved very similarly as the target lines are close to each other. There was no main effect of Group in any time window indicating that one group looked more to the targets than the other. However, the main effect of Target Word was significant in all time windows ($b_{\text{Target Word}} = 43.23$, $p_{\text{MCMC}} < 0.01$), indicating that both groups had a greater target preference for the /ʋ/-targets than for the /b/-targets. The interaction between Target Word and Group was not significant. That is, we did not find an effect of exposure in the segmental reduction condition.

Figure 6. Eye-tracking data for the /b/- and /ʋ/-targets in the test phase of Experiment 1.



In the syllabic reduction condition (Figures 7A and B), the two groups differ in their looking behavior to the *ver*-targets from around 500-800 ms after target onset, whereas they seem to fixate the /f/-targets equally well. The main effect of Group was not significant in either time window. However, the main effect of Target Word was significant from 400ms after target onset onwards ($b_{\text{Target Word}} = 134.25$, $p_{\text{MCMC}} < 0.001$). That is, participants in both groups had a greater target preference for the /f/- than for the *ver*-targets. Importantly, the interaction between Group and Target Word was significant in the time window from 600-800 ms after target onset ($b_{\text{Group x Target Word}} = -56.23$, $p_{\text{MCMC}} < 0.05$). That is, we did find an effect of exposure in the syllabic reduction condition.

Figure 7. Eye-tracking data for the *ver-* and /f/-targets in the test phase of Experiment 1.



Discussion

In Experiment 1, we found that both kinds of reduction, segmental and syllabic, hinder word recognition. The effects of reductions observed in the exposure phase indicate that reduced /b/-words were harder to recognize than canonical /b/-words. Similarly, reduced *ver*-words were harder to recognize than canonical *ver*-words.

In the test phase, we observed an adaptation effect for the syllabic reduction group. The eye-tracking data show that the syllabic reduction group had an advantage over the segmental reduction group when they heard reduced *ver*-words: The syllabic reduction group, who had already been exposed to reduced *ver*-words, showed a greater target preference than the segmental reduction group for whom *ver*-reductions were a new phenomenon. Thus, listeners are able to adapt to syllabic reductions. But we did not observe a similar adaptation effect in the segmental reduction group: Participants in this group did not look more closely to the /b/-targets than participants in the syllabic reduction group. This null effect of learning for the segmental reductions is quite surprising given the learning effect for the syllabic reductions. Intuitively and according to most models of word recognition, it should be easier to adapt to segmental reductions, as they do not deviate as much from canonical pronunciations as syllabic reductions do.

The reason for the null effect in the segmental reduction condition might lie in our materials. To be more precise, the reason might lie in the articulation of our speaker: According to the corpus study we conducted before Experiment 1, a /b/-to-[ʋ]-reduction was most likely to occur after a vowel (40% of the cases) or a nasal (27% of the cases), that is, after a sonorant voiced segment. A closer look at the Experiment 1 stimuli revealed, however, that our speaker systematically devoiced not only all the /v/s but also all the schwas at the end of *woordje*. As the unvoiced schwa is phonetically not a sonorant voiced segment, the /b/s of the following /b/-targets did not appear in an appropriate context for reduction to occur. Experiment 2 was run to address this issue and to replicate the adaptation effect for syllabic reductions.

Experiment 2

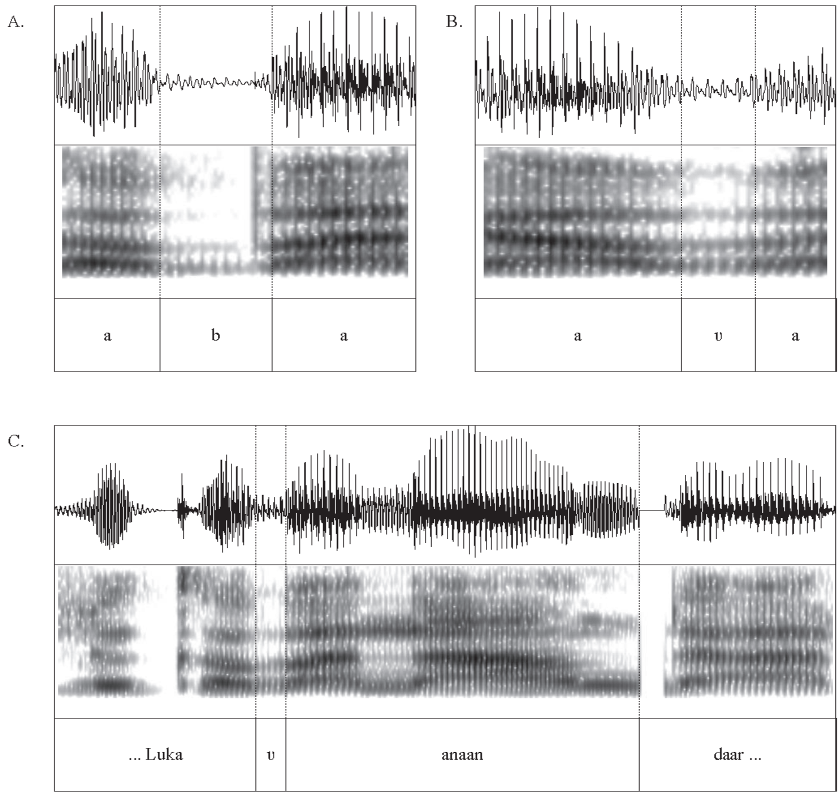
In Experiment 2 we wanted to test whether the null effect for the segmental reduction condition in Experiment 1 was due to an inappropriate reduction context. The same critical target words and design were used, but the words were embedded in a new carrier sentence template: "Nu zegt Luka ... daar links/rechts van het rondje/sterretje/ driehoekje/vierkantje" (literally: 'Now says Luka ... there to the left/right of the circle/star/triangle/rectangle'). The color adjective was omitted to shorten the experiment. Importantly, the beginning of the sentence was changed so that the pre-target word ended in a full vowel (*Luka*); full vowels are less prone to reduction than the pre-target schwa used in Experiment 1. After an open jaw position, as is necessary for /a/, achieving a closure for the next consonant represents an effort for the speaker. Thus, the /b/-targets appeared now not only in an appropriate phonological context (i.e., in intervocalic position) but also in a phonetic context which should make reductions more likely.

Method

Participants. Fifty-two Dutch native speakers were recruited from the Max Planck Institute subject pool and paid for their participation. The participants reported normal hearing and normal or corrected-to-normal vision. None of the participants had taken part in the previous experiment.

Stimulus construction. The new sentences were recorded digitally by a female native speaker of Dutch who also devoiced [v] to [f] naturally. The sentence accent remained on the side disambiguation (*links/rechts*). The recording procedure was virtually identical to the one described in Experiment 1. However, knowing from earlier recordings that this speaker reduces word-internal /b/s naturally to labio-dental approximants, we did not ask her to produce reductions consciously. Instead, we let her additionally record words containing the sequence *a.bV*, where the *a* occurs in the coda of one syllable, the *b* in the onset of the next, unstressed syllable followed by one of the vowels occurring in the original /b/-target words (e.g., *c<u>aba</u>ret*; the underlined sequence corresponded for instance to "… *Luk<u>a ba</u>naan* …" in one of the experimental sentences). Without instructions, the speaker reduced these word-internal [b]s naturally. These naturally reduced [b]s were then extracted and spliced onto the beginnings of the corresponding /b/- and /ʋ/-initial target words (see Figure 8). The pitch of each reduced [b] (a labio-dental approximant) was matched to the surrounding context using the PSOLA algorithm in PRAAT. Intensity was changed where necessary so that the difference in intensity between the reduced [b] and the following vowel corresponded to the difference measured in the source word (where the reduction occurred naturally, e.g., *cabaret*). If there was no Dutch word containing a given sequence *a.bV* word-internally, so that there was no naturally reduced [b] for that vowel, either the reduced [b] preceding a neighboring vowel (in the vowel space) or the reduced [b] preceding schwa was used.

Figure 8. Example for cross-splicing of an experimental sentence containing a /b/-reduction (*banaan* 'banana') in Experiment 2. Figure 8A shows the waveform of a natural *a.ba*-sequence as it occurs in "Luka banaan". Figure 8B shows the waveform of a naturally reduced *a.ba*-sequence as it occurs in *cabaret*. Figure 8C displays the partial waveform of an experimental sentence containing a spliced-in naturally reduced [b] and shows all splicing points.



As the speaker was not asked to produce reductions consciously, we had to create the *ver*-reductions digitally. Therefore, the longest [f] of a set of natural /f/-words sharing the first vowel was selected and lengthened by 20%. This long [f] was then spliced onto all /f/- and *ver*-words containing the corresponding vowel in second position.

Contrary to Experiment 1, every target word had its own token of the first part of the carrier sentence (*Nu zegt Luka*). For the last part of the carrier sentence indicating the position of the geometrical shape (e.g., *daar rechts van het driehoekje*), however, only one version for each combination of position and shape was used. If there was a sudden change in amplitude at the transition between the preceding vowel and the spliced-in reduced [b] or the spliced-in reduced [b] and the following vowel, it was smoothed by adding an amplitude decrease to the corresponding vowel.

Design and procedure. The design was the same as in Experiment 1. The procedure also corresponded to the earlier one, with only small changes. Eye movements of one eye were now recorded with an SR Research EyeLink 1000, sampling at 1 kHz. The right eye was tracked for all participants. As the color adjective was omitted in the carrier sentence, so were the color trials among the fillers. Consequently, a reduced number of 30 fillers were interspersed between the 96 exposure and 96 test trials. An experimental session now took approximately 30 minutes.

Results

Exclusion criteria and statistical testing. Thirty trials (0.3%) were excluded due to imprecise mouse clicks (i.e., not within a circle of 180 pixels around one of the centers of the four words on the screen). In Experiment 2, 100 pixels correspond to 4 cm due to a different screen size; average word length was again 165 pixels. Only trials in which participants clicked on the target word were included in the eye movement analyses. In trials where participants did not fixate the screen, the respective x- and/or y-value was set to the corresponding edge of the screen. Outliers were excluded based on RT criteria following the procedure used in Experiment 1. Thus, in the exposure phase, trials whose residuals were higher than 1500 ms or lower than -1000 ms were removed (13 trials). In the test phase, 10 trials were excluded as their residuals were either below -1200 ms or above 1800 ms. These values were chosen based on visual inspection of a residual plot.

Participants in both groups performed the task accurately. The percentages of correct responses (clicks on the target word) equalled or exceeded 92% in all conditions for both exposure and test phases. Descriptive statistics of RTs for correct trials are displayed in Table 3 for both groups as well as exposure and test phase. This time, participants in the segmental reduction group responded faster than participants in the syllabic reduction group (see also Figures 9 and C1). The overall lower mean RTs in this experiment are probably due to the fact that the female speaker of Experiment 2 spoke faster than the male speaker of Experiment 1. Moreover, the sentences were shorter than in Experiment 1. Both facts led to faster RTs for those subjects that clicked only once the sentence was finished.

The same models and the same procedure as in Experiment 1 were used to analyze the data. Groups and target words were contrast coded in the same way as before.

Table 3. Descriptive statistics of RTs for correct trials for both groups in the exposure and test phases in Experiment 2.

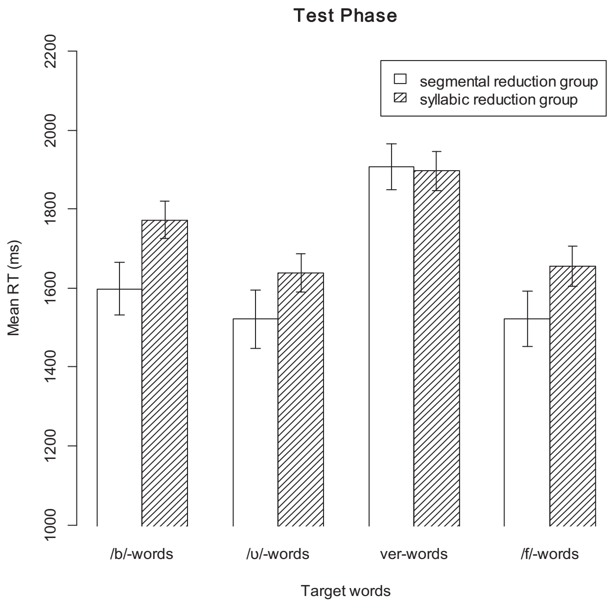
	Segmental reduction group		Syllabic reduction group	
RT in ms	Exposure	Test	Exposure	Test
mean	1493	1641	1633	1753
sd	493	526	426	510
min	500	571	383	458
max	5574	3853	3753	7767

Exposure Phase. Only a summary of the exposure phase results follows. For detailed analyses, see Appendix C. In the RT and eye-tracking data, effects of reduction became evident: The segmental reduction group (who heard reduced /b/-words) showed a smaller target preference for /b/-words than the syllabic reduction group (who heard canonical /b/-words). Vice versa, the syllabic reduction group (who were exposed to reduced *ver*-words) took longer to respond to and showed a smaller target preference for *ver*-words than the segmental reduction group (who heard canonical *ver*-words).

Test phase. Figure 9 displays the mean RTs of the click responses for both groups. The segmental reduction group responded faster than the syllabic reduction group for every word type except for the (reduced) *ver*-words. Compared to the other word types, it is also the *ver*-words that both groups took much longer to react to. Moreover, the syllabic reduction group was also strongly slowed down by the now reduced /b/-words. Statistical analyses showed that in the segmental reduction condition (/b/- and /ʋ/-targets), the main effects of Group and Target Word as well as their interaction were significant ($b_{Group} = 148$, $p_{MCMC} < 0.05$; $b_{Target\ Word} = -111$, $p_{MCMC} < 0.001$; $b_{Group\ x\ Target\ Word} = -64$, $p_{MCMC} < 0.05$). Thus, in responding to reduced /b/- and quasi-canonical /ʋ/-words, the segmental reduction group was significantly faster than the syllabic reduction group. Furthermore, both groups reacted faster to /ʋ/-words than to /b/-words. The effect of exposure (the significant interaction), is driven mostly by the syllabic reduction group who took longer to respond to reduced /b/-words. The main effect of Trial Number was not significant. That is, participants did not get faster in responding to /b/- and /ʋ/-words.

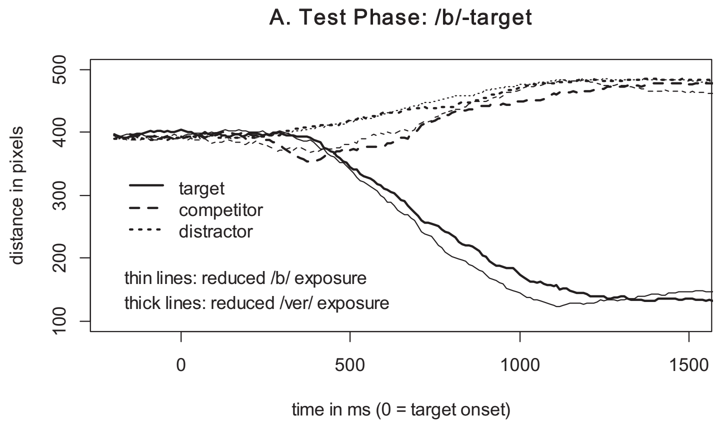
In the syllabic reduction condition (*ver*- and /f/-words), the main effect of Group was not significant. So, no group was faster than the other when reacting to both *ver*- and /f/-words. In contrast, the main effect of Target Word was significant ($b_{\text{Target Word}}$ = -327, p_{MCMC} < 0.001), indicating that both groups responded faster to quasi-canonical /f/- than to reduced *ver*-words. Also the interaction between Group and Target Word was significant ($b_{\text{Group x Target Word}}$ = 143, p_{MCMC} < 0.001). This interaction is driven by the segmental reduction group, as participants in this group were slowed down more strongly by *ver*-reductions which they had not encountered in the exposure phase than the syllabic reduction group (see the four rightmost bars in Figure 9). Finally, also the main effect of Trial Number was significant ($b_{\text{Trial Number}}$ = -1, p_{MCMC} < 0.05), indicating that participants reacted faster to *ver*- and /f/-words during the test phase.

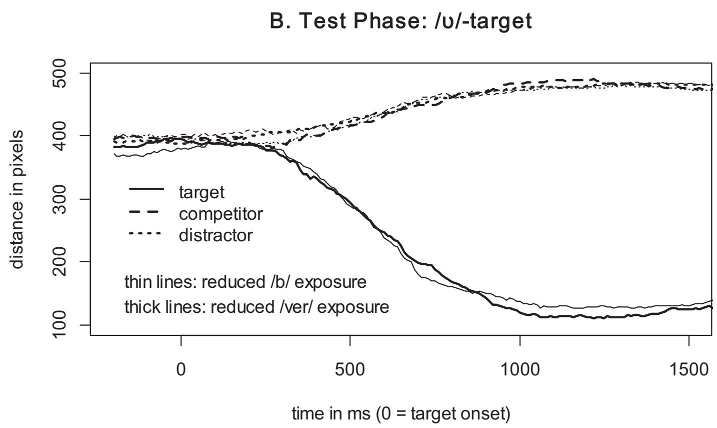
Figure 9. Mean RTs in ms and SEs of both groups for each type of target word in the test phase of Experiment 2. All /b/- and *ver*-words were reduced. All /ʋ/- and /f/-words were presented in quasi-canonical pronunciation.



The corresponding eye-tracking data for the test phase are displayed in Figures 10 and 11. In the segmental reduction condition (/b/- and /ʋ/-words), both groups first fixate the competitor (dashed lines) when hearing a reduced /b/-word, before looking to the target at around 450 ms after target onset (see Figure 10A). Later on, the segmental reduction group (thin lines) looked closer to the /b/-targets. When hearing a quasi-canonical /ʋ/-word, the two groups behaved very similarly (the target lines overlap, see Figure 10B). Statistical analyses did not reveal a main effect of Group. That is, neither of the groups looked more to the target words (when averaged over both conditions) than the other. However, there was a main effect of Target Word in all time windows ($b_{\text{Target Word}} = 60.74$, $p_{\text{MCMC}} < 0.001$) indicating that both groups looked more to the /ʋ/-words than to the /b/-words. The interaction between Group and Target Word was significant from 800-1300 ms after target onset ($b_{\text{Group x Target Word}} = 57.36$, $p_{\text{MCMC}} < 0.001$). Thus, in contrast to Experiment 1, we found an effect of exposure in the segmental reduction condition. The segmental reduction group looked closer to the reduced /b/-targets presumably because they had more experience with /b/-reductions than the syllabic reduction group, who in fact had a small advantage when /ʋ/-words were the targets.

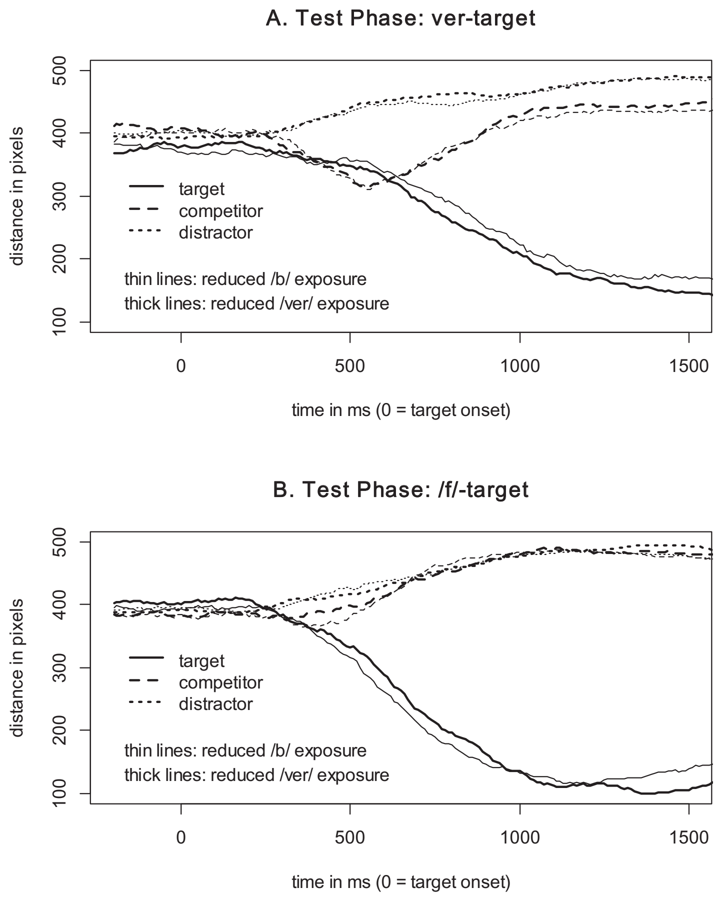
Figure 10. Eye-tracking data for the /b/- and /ʋ/-targets in the test phase of Experiment 2.





In the syllabic reduction condition (*ver*- and /f/-words), both groups looked first to the competitor (dashed lines) when hearing a reduced *ver*-word, but later on, the syllabic reduction group (thick lines) looked closer to the *ver*-targets (see Figure 11A). When an /f/-word was the target, it was the segmental reduction group who looked closer to the targets (see Figure 11B). Statistical analyses did not reveal a main effect of Group. That is, no group looked more to the targets (when averaged over both conditions) than the other. In contrast, there was a main effect of Target Word from 400 ms onwards ($b_{\text{Target Word}}$ = 129.55, p_{MCMC} < 0.001) indicating that both groups had a greater target preference for the /f/-words than for the *ver*-words. The interaction between Group and Target Word was significant in a time window from 700-1000 ms after target onset ($b_{\text{Group x Target Word}}$ = -49.66, p_{MCMC} < 0.05). Thus, we replicated the effect of exposure for the syllabic reduction condition found in Experiment 1.

Figure 11. Eye-tracking data for the *ver*- and /f/-targets in the test phase of Experiment 2.



Discussion

Experiment 2 was conducted to investigate whether the null effect of exposure for the segmental reduction condition found in Experiment 1 was due to an inappropriate phonetic context for the b/-to-[ʋ]-reduction. There were three minor changes between experiments. First, the sentences were shortened by omitting the color indication of the geometrical shapes. Therefore, the number of filler sentences was also reduced. Second, a different speaker, who reduced word-internal intervocalic [b]s naturally to [ʋ]s, was recorded. It seems unlikely that the change of speaker (gender) should matter when it comes to adaptation to a single individual. Third, small changes were made in the way both sets of materials were made. For the segmental reductions, instead of the duration-manipulated stimuli used in Experiment 1, naturally occurring /b/-reductions were spliced onto the /b/- and /ʋ/-stimuli. For the syllabic reductions, instead of the consciously produced *ver*-reductions used in Experiment 1, duration-manipulated material was used to create the *ver*- and /f/-stimuli. The major change in procedure concerned the carrier sentence. In Experiment 1, attention was paid to the fact that the word preceding the targets ended in a vowel (the schwa of *woordje*). However, as the speaker in Experiment 1 devoiced this schwa, a carrier sentence was chosen that ended in a full open vowel (*Luka*… instead of *woordje*…) to ensure that an appropriate phonetic context for the occurrence of /b/-reductions was given.

In the exposure phase, the data show inhibitory effects of reduction on word recognition, similar to the findings from Experiment 1. The target preference for /b/-words was larger for the syllabic reduction group, who heard them in the canonical form, than for the segmental reduction group, who heard them in the reduced form. The same is observed for the *ver*-words. Here, the segmental reduction group, who heard these words in the canonical form, showed a larger target preference than the syllabic reduction group, who heard these words in the reduced form. These reduction costs are also reflected in the RT data. That is, the new implementation of the reductions again hindered word recognition.

In the test phase, we again observed an adaptation effect for the syllabic reduction condition: The interaction between Group and Target Word, which reflects an effect of exposure, was significant not only in the eye-tracking data but also in the RT data. That is, the syllabic reduction group could make use of their prior exposure to reduced *ver*-words and recognized new reduced *ver*-words faster and had a greater target preference than the segmental reduction group for whom the *ver*-reductions were new. Thus, we replicated the adaptation effect for syllabic reductions found in Experiment 1.

Importantly, we also observed an adaptation effect for the segmental reduction condition: The analyses of the RT and the eye-tracking data revealed an effect of exposure. That is, the segmental reduction group, who had already heard /b/-reductions in the exposure phase, responded faster to new reduced /b/-words and showed a greater target preference than the syllabic reduction group for whom the /b/-reductions were a new phenomenon. This effect was larger and stretched out over a longer period of time than the replicated adaptation effect for syllabic reductions. The fact that a change in the carrier sentence (i.e., the word preceding the reduced /b/-target ending in a full vowel instead of a devoiced schwa) led to the observation of an adaptation effect for /b/-reductions shows how context-sensitive listeners are.

Figures 10A and 11A show that both groups first look to the competitor (dashed lines) when hearing reduced /b/- and *ver*-words. This suggests that listeners first interpret the incoming speech according to the acoustical information they get as [ʋ] or [f] respectively. Only later, they re-evaluate their categorization as the rest of the word becomes available (for the untrained group) or learning intervenes (for the trained group).

General Discussion

Previous studies investigating mechanisms compensating for reduced speech have focused on language- or speech-dependent characteristics. The current study is the first to investigate a compensation mechanism which takes the speaker into account. Reduced pronunciation variants are a phenomenon of stylistic variation that depend on individual speakers but are not inherent to them. That is, it is the speaker's choice whether he or she reduces certain words or parts thereof (Mitterer & Ernestus, 2006). In the exposure phase, each group of participants heard the speaker produce a certain type of reduction (either segmental, i.e., /b/ → [ʋ], or syllabic, i.e., *ver*- → [f], reductions). In the test phase then, the speaker produced two types of reduction. In this test phase, the syllabic reduction group was better than the segmental reduction group in recognizing new reduced *ver*-words. This effect was found in both experiments. In Experiment 1, we did not find a similar advantage for the segmental reduction group. Participants in this group were not better at recognizing new reduced /b/-words than the syllabic reduction group who had not previously been exposed to /b/-reductions. Changing the carrier sentence, so that the phonological and phonetic context made /b/-reductions more likely, solved this problem. In Experiment 2, where these changes were applied, an adaptation effect was found for the segmental reduction group. Participants in this group were now better at recognizing new reduced /b/-words than the syllabic reduction group. Listeners are thus able to adapt to specific reduction styles of a speaker and this helps them to compensate for speech that has been reduced in a way consistent with prior exposure.

The current study thus extends experimental evidence of perceptual learning for naturally-occurring variation. Mitterer and McQueen (2009a), for example, showed that listeners can adapt to an unfamiliar regional accent in a second language. They exposed Dutch participants to either Scottish or Australian English videos with Dutch, English or no subtitles. The subsequent test phase showed that the lexical information provided by the English subtitles enhanced adaptation effects, whereas the information provided by the Dutch subtitles, which was inconsistent with the spoken English word forms, hindered adaptation. It is unclear, however, which specific phenomena the listeners tuned in to. Dahan et al. (2008) showed that listeners can tune in to regional variation of a given speaker's native language. In their study, listeners adapted to a regional phenomenon (the raising of the vowel /æ/ before /g/, but not before /k/) which facilitates word recognition by reducing ambiguity in the initial part of words like *bag* and *back*. The current study is the first to show a similar adaptation effect for stylistic variation. Listeners attuned to segmental and syllabic reductions which, importantly, increased ambiguity between existing words (e.g., *baron* pronounced as [ʋarɔn] can easily be confused with the Dutch word [ʋarɔm] *waarom* 'why'). This increased ambiguity was found not only for extreme cases such as the quasi-minimal pairs used in the test phase, but also in the exposure phase where reduced forms hindered word recognition even in target-"competitor" pairs of hardly overlapping words (e.g., the target *bikini* pronounced as [ʋikini] and canonical [ʋustənɛɪ] *woestenij* as "competitor"). Hence, the current study is also the first to show that perceptual learning not only applies to artificially constructed ambiguous sounds (e.g., Maye, et al., 2008; McQueen, et al., 2006; Norris, et al., 2003), but also to ambiguities occurring in natural continuous speech.

Importantly, perceptual learning about segmental and syllabic reductions was applied to previously unheard words in the test phase. Listeners could use the experience they gained with one type of reduction in the exposure phase to recognize new reduced words of the same reduction type in the test phase. McQueen et al. (2006) observed generalization of learning about ambiguous fricatives ([$^{s}/_{f}$]) to new words. They argued that, for learning about a sublexical unit to be applied to previously unheard words, an abstraction process concerning those units must take place at a prelexical level. Mitterer et al. (2011) conducted a similar perceptual learning experiment about an ambiguous Mandarin tone contour and tested the effect not only for new but also for repeated words. The learning effect was slightly larger for previously heard than for previously unheard words. This suggests that there are small additional effects of episodic learning compared to large effects of prelexical abstraction.

Based on the finding that learning about reductions generalized to new words, we would expect to observe the adaptation effect early in the eye-tracking data. However, we observed quite late effects, starting at around 600-800 ms after target onset. In contrast, competitor preferences, although not significant, were observed much earlier (as early as 200 ms - 250 ms after target onset). This seems to suggest that the bottom-up information in the speech signal has a strong influence on speech recognition (e.g., Marslen-Wilson, 1987) and learning needs time to override the signal-driven interpretation (i.e., “what-you-hear-is-what-you-get”).

However, Mitterer and Reinisch (in press), who investigated the time course of perceptual learning about ambiguous fricatives, found early learning effects in their eye-tracking data. They concluded that perceptual learning is indeed perceptual rather than post-perceptual once it is complete. Mitterer and Reinisch (in press) observed these early learning effects using an eye-tracking paradigm with a classical exposure-test design. It thus seems unlikely that the late effects found in the present study are caused by the combination of a printed-word eye-tracking task and a learning paradigm. Furthermore, the participants in the study by Mitterer and Reinisch (in press) had to make conscious metalinguistic judgments about the speech sounds they heard by clicking on printed words whose spoken realizations deviated from the canonical pronunciations. As this task did not affect the timing of the effects in Mitterer and Reinisch (in press), it seems unlikely that it should have done so in the experiments presented here.

If perceptual learning should and does in general occur early, and methodological issues have been ruled out as possible causes, what might then be the reason for the late learning effects observed in this study? One point in which the present study differs from the study by Mitterer and Reinisch (in press) and other classical perceptual learning studies (e.g., McQueen, et al., 2006) regards the mapping rules listeners must generate or adjust to correctly recognize the deviant sound(s). In the classical perceptual learning paradigm, listeners have to learn to include a previously unknown ambiguous sound into one of two possible phoneme categories. They must thus generate a new mapping rule for that sound, and that sound is supposed to be truly ambiguous. Ideally, it is judged as belonging to one of the two categories in 50% of the cases. Apart from the pretested [$^{b}/_{ʋ}$] sounds in Experiment 1 (for which no learning was found), none of the reduced sounds used in this study were truly ambiguous. The naturally occurring [$^{b}/_{ʋ}$]s used in Experiment 2 as the replacement for /b/ were likely more /ʋ/-like than /b/-like. The long [fː]s replacing *ver*-s in Experiments 1 and 2 were clearly more /f/-like than *ver*-like. Listeners thus learned that a particular speaker was likely to pronounce a /b/ as an [ʋ] and hence that an existing sound ([ʋ]) mapped onto two categories for that speaker (i.e., /ʋ/ and /b/). Their perception of an [ʋ] might therefore have shifted from judging it as /ʋ/ in most cases to judging it as /ʋ/ in 80% and as /b/ in 20% of the cases. With this kind of learning, the initial signal-driven hypotheses still strongly favor the canonical form, and only when later-arriving segments rule that form out can the learning take effect. Therefore, as soon as listeners receive evidence that a particular sound can map onto more than one category, the learning process likely needs more time to take effect.

The generalization of learning about reductions across words is evidence for an abstractionist mode of lexical access and therefore also for abstractionist models of the mental lexicon. It suggests that compensation mechanisms for both segmental and syllabic reductions involve a phonologically abstract prelexical level. In the case of segmental reduction, prelexical mapping from an ambiguous reduced [b] to the segmental representation of a /b/ should work quite similarly to the mapping of ambiguous [$^{s}/_{f}$] to either /s/ or /f/. However, the mapping from a reduced *ver*- (i.e., from the single segment [fː] in our case) to the syllable *ver*- requires a larger prelexical representation than is specified in current abstractionist models (e.g., Gaskell & Marslen-Wilson, 1997; McClelland & Elman, 1986; Norris & McQueen, 2008). One can think of several possibilities regarding the nature of this prelexical representation. First of all, considering the morphological status of *ver*-, this representation might be morphemic. This would be in line with models of word recognition which assume that morphemes are segmented at the prelexical level (Taft, Hambly, & Kinoshita, 1986). Prelexical morphemic representations are a necessary consequence of prelexical morphological parsing. The input [fː] could thus be mapped onto a representation of the morpheme *ver*-.

Another possibility, which in the case of *ver*- cannot be teased apart from the morphemic account, is that prelexical representations are syllabic. This is certainly a more extreme claim as then every syllable (and there are approximately 21,800 syllables in the Dutch language as calculated on the basis of the CELEX database [Baayen, Piepenbrock, & van Rijn, 1993]) would need its own representation.

A third option assumes that there is a more general prelexical process through which frequently-occurring sequences of segments are bound together. Prelexical processes are unaffected by knowledge stored in the mental lexicon, but may be sensitive to transitional probabilities between speech sounds (Pitt & McQueen, 1998). In a series of experiments, Vitevitch and Luce (1999) dissociated lexical and prelexical effects in spoken word recognition by investigating probabilistic phonotactics and neighborhood activation. Their findings support the idea that the prelexical level is sensitive to sequential probabilities. In our case, *ver*- has a very high transitional probability. Prelexical mechanisms would thus have to capture the fact that such high probability sequences can also be realized in a reduced form.

Finally, the prelexical representation might be a combination of segmental representations and an abstract representation of the syllable as a prosodic unit. During speech decoding, a Prosody Analyzer (Cho, McQueen, & Cox, 2007) may use the supra-segmental information available in the speech stream to compute the prosodic structure of the utterance. This structure serves then to unify sequences of segments. At the lowest level in the prosodic hierarchy, the syllable is bound to a combination of segments. In this process, it might be possible to match only one segment ([fː] in the case of *ver*-) to a syllable. Note that such mappings are already necessary for syllabic sonorants, such as in the word bottle [bɔːtl̩], in which the final [l̩] is syllabic. Dutch listeners are also confronted with syllabic consonants. In Groningen, for example, speakers may pronounce a verb like *laten* /latən/ 'to let' as [latn̩]. Future research is required to disentangle these four possibilities concerning what we have called syllabic reductions. The current results, however, already suggest that prelexical representations or processes are required that go beyond those assumed in current abstractionist models.

Our data show that there is abstraction in spoken-word recognition, and hence suggest that there is more to this process than the storage of episodes of actual pronunciations. Variation phenomena such as reductions are often the raison d'être for episodic models (Hawkins, 2003). However, the present study suggests that naturally-occurring systematic variation, at least the segmental and syllabic reductions we examined, is dealt with by a process of phonological abstraction. Cutler et al. (2010) ran simulations with a purely episodic model (MINERVA-2; Hintzman, 1986) to test whether an exemplar-based theory could account for lexical generalization of perceptual learning about ambiguous fricatives. That was not the case. The current study presents similar kinds of learning processes. For the segmental reductions, the reduced element is a sound in between a bilabial voiced stop and a labio-dental approximant. Learning for this ambiguous sound is likely to be very similar to the ambiguous sound [$^{s}/_{f}$] (between /s/ and /f/) studied by Cutler et al. It is thus unlikely that a purely episodic model could account for the lexical generalization for reduced segments found here, and thus also for the generalization about reduced syllables. Our data thus speak against any model of the mental lexicon that does not provide a mechanism for generalization of learning to previously unheard words.

Consider also, in contrast to classical abstractionist models, the variant proposed by Ranbom and Connine (2007): Their frequency-based account contains gradient and multiple phonological representations for each entry in the lexicon. Citation forms and pronunciation variants of words are stored, graded in strength according to experienced frequency. Thus, no special processes are required to compensate for variation such as reduction in the speech signal. This model, however, has problems accounting for our findings, as there is no mechanism for generalization.

Another abstractionist model that has difficulties dealing with our data is the featurally underspecified lexicon (FUL) model (Lahiri & Reetz, 2002, 2010). In this model, the entries in the mental lexicon are based on phonological features. The first segment of the Dutch word *baron*, for example, is specified in the lexicon as [labial] and [voice]. This segment is not explicitly specified as [plosive], as stops are assumed to be the default for manner of articulation. The features that are extracted for the first segment of the reduced speech input [ʋarɔn] are categorized as [labial], [continuant] and [voice]. The extracted feature [continuant] does not mismatch the underspecified representation in the lexicon. That is, according to the FUL model, it does not matter that /b/s are reduced to [ʋ]s. Therefore, there is no need for listeners to adapt to these /b/-reductions and adaptation should thus not occur. The fact that we find an adaptation effect for /b/-reductions contradicts the assumption of the FUL model.

Our data thus support fully specified abstractionist models which provide a mechanism for generalization and which include a representation of sequences larger than segments at the prelexical level. It is important to emphasize here, however, that abstractionist models relying completely on prelexical compensation for reductions cannot be the full story. Listeners will often encounter reductions, and especially reduced prefixes, from formerly unknown speakers. These need to be recognized as well, and there is evidence that storage of variant forms can contribute to that. Mitterer and Russell (2013) found that processing costs for reduced prefixes occur mainly for low-frequency words. This can easily be explained by an episodic model, because prefix reductions occur less often in low-frequency words (Pluymaekers, et al., 2005). Such episodic storage, however, does not make prelexical reconstruction superfluous. It has been repeatedly found that, compared to reduced forms, canonical forms have an advantage in speech comprehension (e.g., Ernestus, 2009; Pitt, Dilley, & Tat, 2011). Adapting to a given speaker's reduction style allows listeners to access canonical forms, which apparently facilitates recognition.

Taken together, the available data are therefore consistent with a hybrid model of speech perception, one that includes both abstractionist and episodic modes of processing, and inconsistent with either theoretical extreme. Prelexical abstraction offers an explanation for generalization of learning to previously unheard words, and hence a plausible account of how listeners recover from specific styles of reduction.

Conclusions

The present study provides the first evidence for prelexical perceptual learning about reductions, a phenomenon of stylistic variation which is common in everyday casual speech. We have shown that adapting to reduced forms involves more than just storing episodes of those forms, because learning generalized to new words. We have therefore argued that abstract prelexical representations mediate the mapping of reduced forms onto the lexicon, and that these representations are the locus of the perceptual learning effect. We have also argued that models of spoken-word recognition need to be revised with respect to the nature of these representations. The demonstration that learning about syllabic reductions generalized to new words requires prelexical multi-segment representations.

Appendix A: Critical Test Stimuli

Table A1. Target and competitor words of the test phase for the syllabic and segmental reduction condition with their word frequency per million according to SUBTLEX-NL (Keuleers, Brysbaert, & New, 2010).

Test phase: segmental reductions					
Target	English Translation	Word Freq.	Competitor	English translation	Word Freq.
baatzuchtig	selfish	0.0	waanzinnig	insane	8.3
bagage	luggage	17.9	waartegen	against what	2.7
balans	balance	9.7	waarheen	where to	31.7
banaan	banana	5.3	waarnaast	beside which	0.0
barak	barrack	3.5	waartoe	for which	4.0
baron	baron	7.7	waarom	why	2141.7
basaal	basic	0.0	waarna	whereupon	3.3
bavianen	baboons	0.9	waardeloosheid	worthlessness	0.0
balkondeur	balcony door	0.0	waxinekaars	tea light	0.0
baldadig	wanton	0.0	walhalla	Walhalla	0.7
Berlijn	Berlin	16.8	welaan	well then	0.0
bermuda	bermuda shorts	1.9	welzeker	certainly	0.0
benzine	petrol	24.7	wellevend	courteous	0.0
berinnen	female bears	0.0	weerspiegelen	to reflect	0.3
binnendoorweg	short cut	0.0	wisselvallig	changeable	0.4
binnenskamers	indoors	0.4	wispelturig	fickle	1.1
biljoenen	trillions	0.6	wilskrachtig	strong-minded	0.0
biljetten	bills	3.7	Wilhelmus	Wilhelmus	0.0
biscuitblik	biscuit tin	0.0	wiskundig	mathematical	1.1
bijvoeglijk	adjectival	0.3	wijdmazig	wide-meshed	0.0
bijdehand	quick-witted	4.8	weinigzeggend	uninformative	0.0
boosaardigheid	malice	0.6	woonwerk-verkeer	commuter traffic	0.0
boerenjongen	country boy	0.6	woensdagavond	Wednesday evening	1.4
boerenkool	kale	0.0	woestenij	wasteland	1.6
Average		4.1	Average		91.6

Test phase: syllabic reductions					
Target	English Translation	Word Freq.	Competitor	English translation	Word Freq.
verleuteren	waste one's time talking	0.0	vleugel	wing	8.8
verlustigen	to amuse	0.0	vluchtig	cursory	1.6
verlangzamen	to slow down	0.0	fladderen	to flutter	0.6
verlaging	reduction	0.0	vlagen	bursts	0.7
verlangen	(to) desire	15.7	flanken	flanks	1.1
verlelijken	to uglify	0.0	vleselijk	carnal	0.3
verlening	grant	0.0	vlegel	flail	0.8
verleppen	to wither	0.0	flatje	little flat	1.4
verliefd	in love	107.2	vlieg	fly	29.1
verlies	loss	49.1	vlies	fleece	1.9
verliezer	loser	6.5	vlieger	kite	3.8
verlichting	lighting	4.3	vlinders	butterflies	3.2
verleider	seducer	0.6	vleien	to flatter	1.5
verloofd	engaged	14.4	vloot	fleet	13.2
verlof	leave	11.0	vlot	raft	6.4
verlokkelijk	tempting	0.0	vlokken	flakes	0.3
verlossen	to release from	4.5	flossen	to floss	0.8
verloedering	degradation	0.4	vloeibaar	liquid	2.4
verregenen	to rain off	0.0	vreemdeling	stranger	10.5
verrekenen	to settle	0.0	vreselijk	terrible	112.4
verrekken	to strain	1.6	vrekken	misers	0.0
verrichten	to perform	4.8	frictie	friction	0.0
verruilen	to swap	0.0	fruit	fruit	12.9
verroeren	to stir	0.7	vroeger	formerly	124.4
Average		9.2	Average		14.1

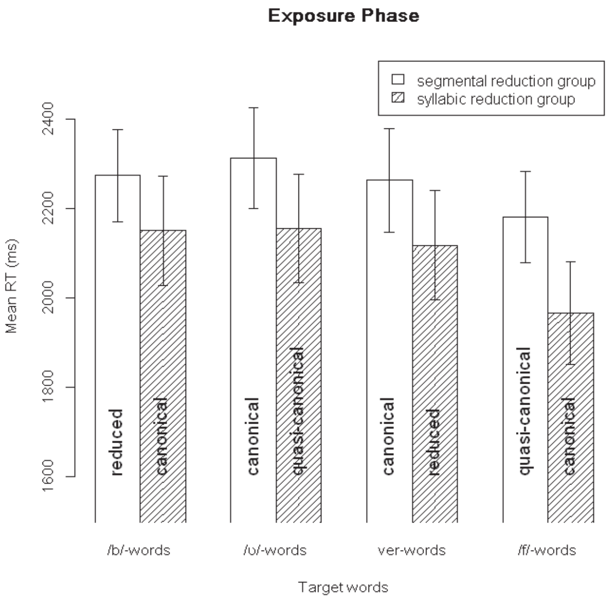
Appendix B: Detailed Analysis of the Exposure Phase Results of Experiment 1

Figure B1 shows the mean RTs in the exposure phase for both groups and each type of target word sampled in two conditions (segmental and syllabic reduction condition). The syllabic reduction group responded faster than the segmental reduction group in both conditions. There was, however, no main effect of Group. In contrast, the main effect of Trial Number was significant not only in the segmental reduction condition ($b_{\text{Trial Number}} = -5$, $p_{\text{MCMC}} < 0.001$) but also in the syllabic reduction condition ($b_{\text{Trial Number}} = -5$, $p_{\text{MCMC}} < 0.001$) indicating that participants reacted faster over the course of the exposure phase.

For the segmental reduction condition, Figure B1 shows that the syllabic reduction group responded to both /b/- and /ʋ/-words equally fast, whereas the segmental reduction group was somewhat faster to respond to (reduced) /b/- than to (canonical) /ʋ/-words. In this case, the main effect of Target Word was not significant. That is, participants, irrespective of group, did not react faster to either /b/- or /ʋ/-words. Also the interaction between Target Word and Group was not significant. Thus, we did not observe an effect of reduction here: The segmental reduction group did not take longer to respond to reduced /b/-words than the syllabic reduction group who heard them in their canonical form.

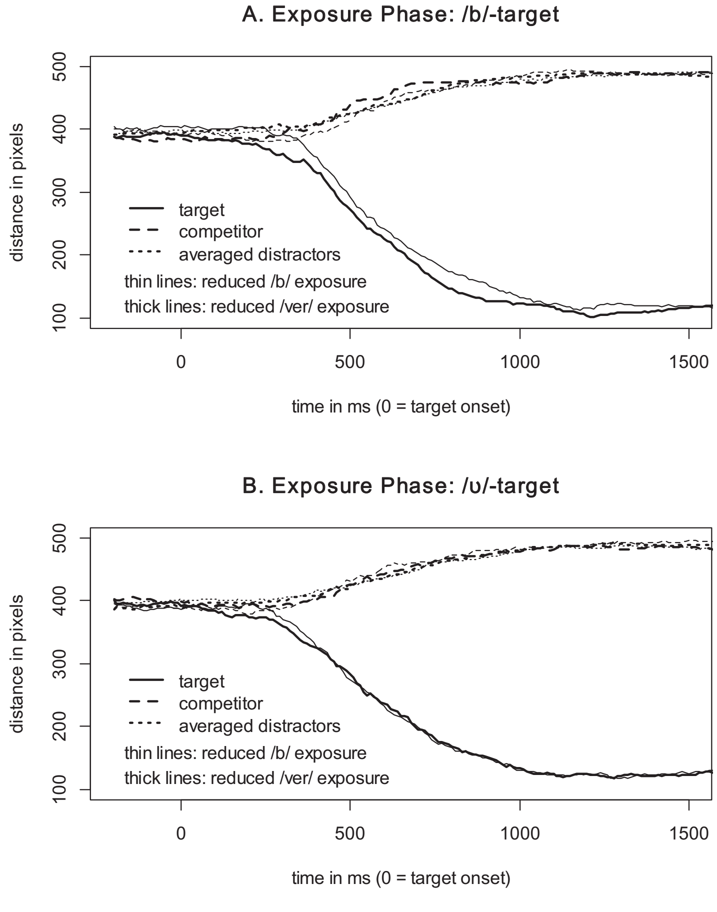
In the syllabic reduction condition, both groups needed more time to respond to *ver*- than to /f/-words, but this difference was bigger for the syllabic reduction group: The main effect of Target Word was significant ($b_{\text{Target Word}} = -113$, $p_{\text{MCMC}} < 0.001$). Note that, because the *ver*-targets were mapped onto -0.5, a negative *b*-value indicates that both groups took longer to respond to the *ver*-targets (the multiplication of a negative *b*-value with -0.5 results in a positive estimate which needs to be added to the intercept). This main effect was moderated by an interaction between Target Word and Group which proved to be marginally significant ($b_{\text{Group x Target Word}} = -71.35$, $p_{\text{MCMC}} = 0.057$). That is, we found a weak effect of reduction, as the syllabic reduction group was slowed down more strongly by the *ver*-words, which they heard in their reduced form, than the segmental reduction group, who heard them in their canonical form.

Figure B1. Mean RTs in ms and SEs of both groups for each type of target word in the exposure phase of Experiment 1. The term 'quasi-canonical' indicates that the stimuli of the control conditions were slightly manipulated: The [ʋ]s of /ʋ/-words were replaced by [$^{b}/_{ʋ}$]s and the [f]s of /f/-words were replaced by [fː]s (for more details see main text).



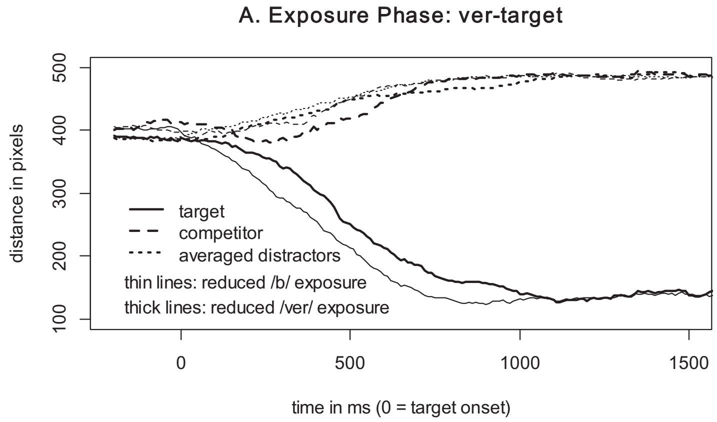
The corresponding eye-tracking data for the exposure phase are displayed in Figures B2 and B3. Figures B2A and B2B show the results for the /b/- and /ʋ/-target words (segmental reduction condition). While the mean fixation of the syllabic reduction group (represented by the thick lines) is closer to the /b/-words than the mean fixation of the segmental reduction group (represented by the thin lines, see Figure B2A), the two groups do not differ in their looking behavior to the /ʋ/-words (the two target lines overlap, see Figure B2B). Statistical analyses showed that there was no main effect of Target Word in any time window (which would have indicated that both groups preferred to look to one target word over the other). However, the main effect of Group was significant in an early time window from 200-400 ms after target onset ($b_{\text{Group}} = 28.91$, $p_{\text{MCMC}} < 0.05$). The positive *b*-value indicates that the syllabic reduction group (mapped onto 0.5) showed a greater target preference for the /b/- and /ʋ/-targets than the segmental reduction group. The interaction between Group and Target Word was significant from 400-900 ms after target onset ($b_{Group\ x\ Target\ Word} = -41.02$, $p_{MCMC} < 0.05$). Thus, there was an effect of reduction in the segmental exposure phase, as participants in the segmental reduction group (hearing reduced versions; thin lines) did not look as closely to the /b/-words as the syllabic reduction group (hearing canonical versions; thick lines), while both groups behaved in the same way for the /ʋ/-targets.

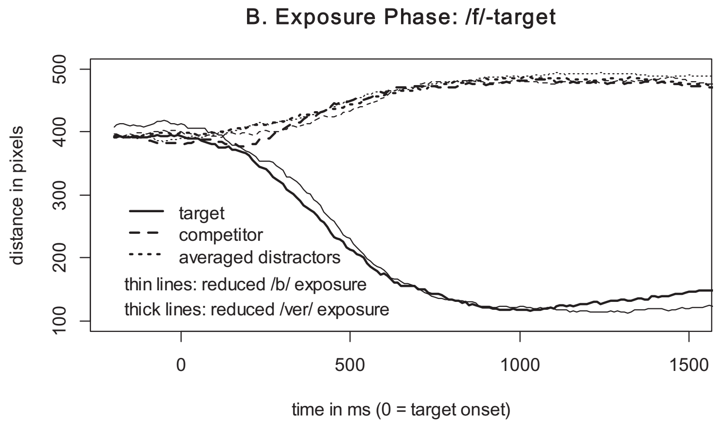
Figure B2. Eye-tracking data for the /b/- and /ʋ/-targets in the exposure phase of Experiment 1.



The eye-tracking results for the syllabic reduction condition of the exposure phase are shown in Figures B3A and B3B. There is a clear difference in the looking behavior of the two groups when the target word is a *ver*-word (see the distance between the target lines in Figure B3A). The segmental reduction group (thin lines) looked closer to the targets than the syllabic reduction group (thick lines) did. For the /f/-words, both groups behaved very similarly, with the segmental reduction group looking slightly less to the targets in the first 500 ms after target onset. Running the linear mixed-effects model did not reveal a main effect of Target Word, but a significant main effect of Group was found in the early time window from 200-400 ms after target onset (b_{Group} = -26.68, $p_{\text{MCMC}} < 0.05$). As the *b*-value is negative, it is the segmental reduction group (mapped onto -0.5) that showed a greater target preference for the *ver*- and /f/-words than the syllabic reduction group. The interaction between Group and Target Word was significant in a time window from 200-800 ms ($b_{\text{Group x Target Word}}$ = 75.96, $p_{\text{MCMC}} < 0.001$). So we also found an effect of reduction in the syllabic exposure phase, as participants in the segmental reduction group (thin lines) looked more to the *ver*-words (hearing them in their canonical form) than participants in the syllabic reduction group (thick lines) who were exposed to reduced versions. The two groups did not differ much when looking to the /f/-words.

Figure B3. Eye-tracking data for the *ver*- and /f/-targets in the exposure phase of Experiment 1.

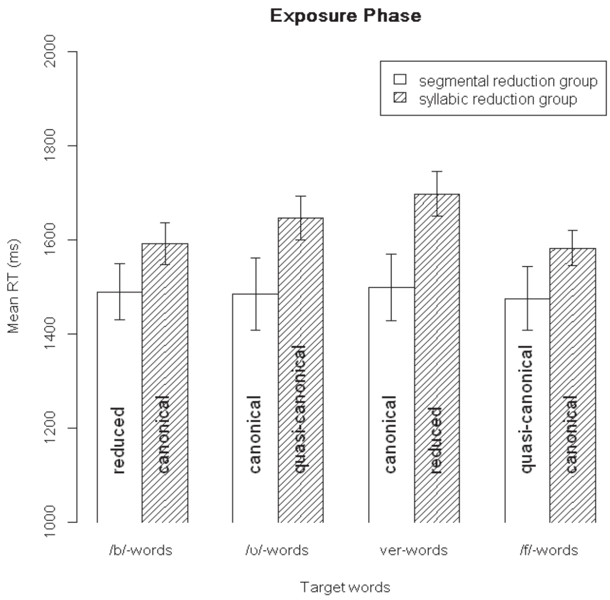




Appendix C: Detailed Analysis of the Exposure Phase Results of Experiment 2

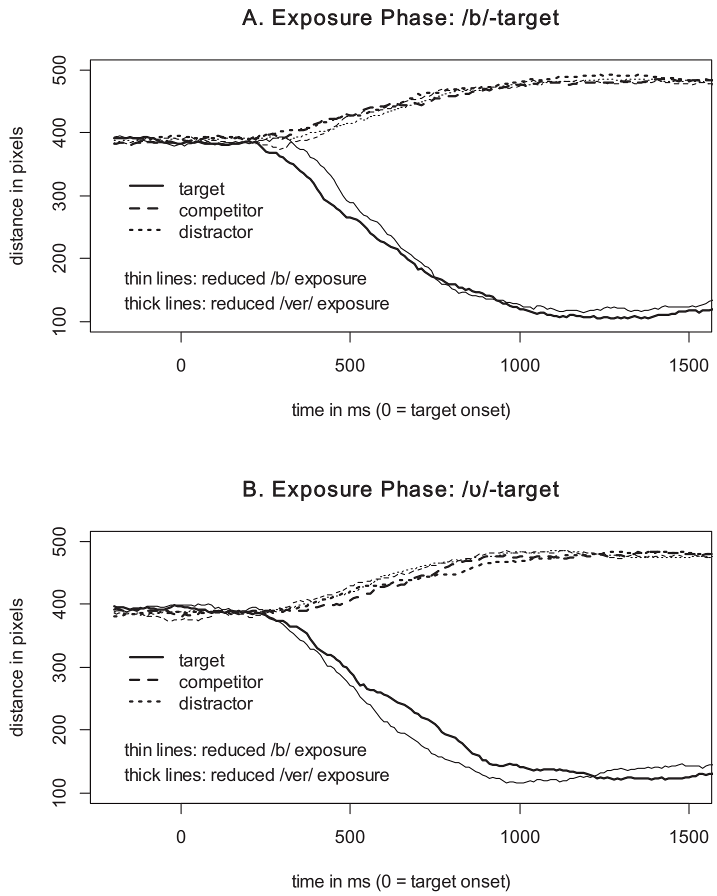
The mean RTs of the click responses for both groups sampled in two conditions are shown in Figure C1. As can be seen, the segmental reduction group responded numerically faster than the syllabic reduction group for every word type. While the segmental reduction group responded to every word type equally fast, the syllabic reduction group seems to be slowed down by quasi-canonical /ʋ/- and reduced *ver*-words (see higher bars for these words). Statistical analyses revealed a significant main effect of Group in both conditions ($b_{\text{Group}} = 132$, $p_{\text{MCMC}} < 0.05$ for the segmental reduction condition and $b_{\text{Group}} = 153$, $p_{\text{MCMC}} < 0.05$ for the syllabic reduction condition). That is, the syllabic reduction group responded overall more slowly than the segmental reduction group (the syllabic reduction group was mapped on 0.5 in the contrast coding). The main effect of Trial Number was significant in both conditions ($b_{\text{Trial Number}} = -2$, $p_{\text{MCMC}} < 0.001$ for the segmental and syllabic reduction condition), indicating that participants reacted faster over the course of the exposure phase. In the segmental reduction condition (/b/- and /ʋ/-words), the main effect of Target Word was not significant. That is, both groups did not respond faster to one word than to the other. In contrast, the interaction between Group and Target Word was significant ($b_{\text{Group x Target Word}} = 58$, $p_{\text{MCMC}} < 0.05$). This effect was driven mostly by the syllabic reduction group, who took longer to respond to quasi-canonical /ʋ/-words. In the syllabic reduction condition (*ver*- and f/-words), the main effect of Target Word as well as the interaction between Group and Target Word was significant ($b_{\text{Target Word}} = -72$, $p_{\text{MCMC}} < 0.001$ and $b_{\text{Group x Target Word}} = -92$, $p_{\text{MCMC}} < 0.001$). The main effect of Target Word indicates that both groups responded faster to /f/-words than to *ver*-words. The interaction reflects an effect of reduction, driven by the syllabic reduction group who was strongly slowed down by the reduced *ver*-words.

Figure C1. Mean RTs in ms and SEs of both groups for each type of target word in the exposure phase of Experiment 2. The term 'quasi-canonical' indicates that the stimuli of the control conditions were slightly manipulated: The [ʋ]s of /ʋ/-words were replaced by naturally occurring /b/-reductions and the [f]s of /f/-words were replaced by [fː]s.



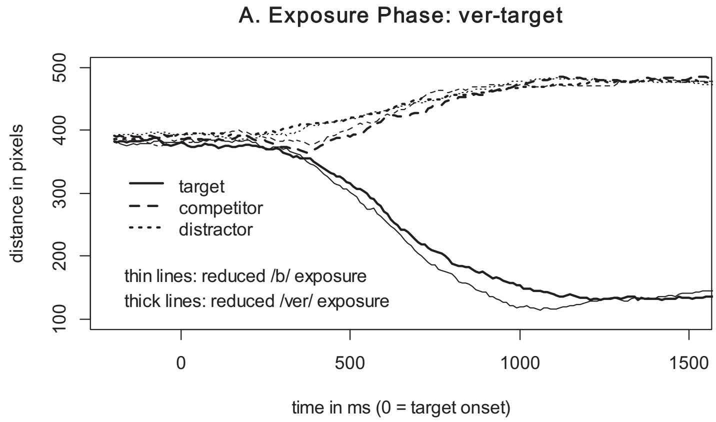
The corresponding eye-tracking data for the exposure phase are displayed in Figures C2 and C3. Figures C2A and C2B show the results for the /b/- and /ʋ/-targets, that is, the segmental reduction condition. It can be seen that for the /b/-words the target lines representing the two groups (thin lines for the segmental reduction group, thick lines for the syllabic reduction group) differ early on (from around 300 ms after target onset, see Figure C2A), whereas they differ somewhat later (from around 500 ms after target onset, see Figure C2B) for the /ʋ/-words. Importantly, the position of the lines switched: For /b/-targets, the syllabic reduction group (hearing canonical /b/-words) looked closer to the target, whereas for the /ʋ/-words, the segmental reduction group showed a greater target preference (hearing canonical /ʋ/-words). Statistical analyses showed that the main effect of Target Word was significant in a small time window from 1300-1500 ms after target onset ($b_{\text{Target Word}} = -19.56$, $p_{\text{MCMC}} < 0.05$). In this time window, participants in both groups had a preference for the /b/-targets over the /ʋ/-targets (as the /b/-targets are mapped onto -0.5 and the b-value is negative). The main effect of Group was not significant in either time window. That is, no group had a greater target preference than the other. The interaction between Group and Target Word, however, was significant from 300-800 ms after target onset ($b_{\text{Group x Target Word}}$ = -69.83, p_{MCMC} < 0.001). Thus, we found an effect of reduction that went into the expected direction: The segmental reduction group who was exposed to reduced /b/-words did not look as closely to the /b/-targets as the syllabic reduction group. The opposite was true for the /ʋ/-targets.

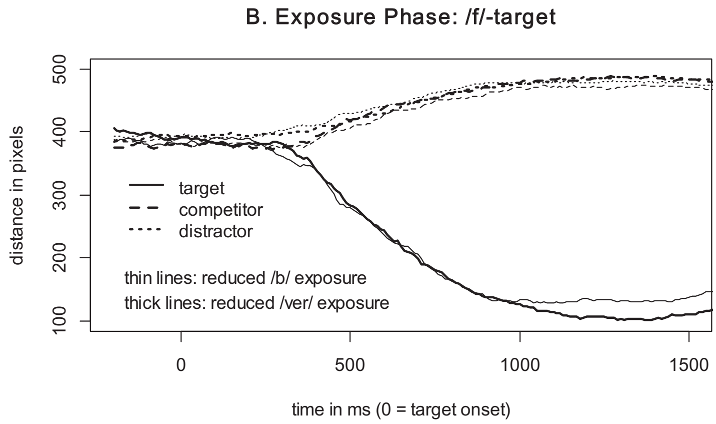
Figure C2. Eye-tracking data for the /b/- and /ʋ/-targets in the exposure phase of Experiment 2.



The eye-tracking results for the *ver*- and /f/-targets (syllabic reduction condition) are shown in Figures C3A and C3B. The target lines differ quite late from around 600-1200 ms after target onset for the *ver*-words (see Figure C3A) and from around 1000 ms onwards for the /f/-words (see Figure C3B). However, the differences in mean fixations of the two groups correspond again to expectations: The segmental reduction group (represented by the thin lines and hearing canonical *ver*-words) looked closer to the *ver*-targets than the syllabic reduction group did (hearing reduced *ver*-words). For the /f/-targets, it is the syllabic reduction group (hearing canonical /f/-words) who had a greater target preference than the segmental reduction group (hearing quasi-canonical /f/-words with a long [fː] at the beginning). Statistical analyses did not reveal a main effect of Group in either time window. That is, the groups did not differ in their target preference for *ver*- and /f/-targets respectively. The main effect of Target Word, however, was significant in the time window from 400-800 ms after target onset ($b_{\text{Target Word}} = 30.01$, $p_{\text{MCMC}} < 0.01$). This indicates that both groups looked closer to the /f/-targets than to the *ver*-targets (as the *b*-value is positive and the /f/-targets are mapped onto 0.5). Also the interaction of Group and Target Word was significant in a time window from 700-1400 ms after target onset ($b_{\text{Group x Target Word}} = 44.65$, $p_{\text{MCMC}} < 0.01$). Thus, also in the syllabic reduction condition, we found an effect of reduction going in the expected direction with closer looks to the *ver*-words if they were heard in unreduced form by the segmental reduction group.

Figure C3. Eye-tracking data for the *ver*- and /f/-targets in the exposure phase of Experiment 2.





Use what you can: Storage, abstraction processes and perceptual adjustments help listeners recognize reduced forms

Chapter 4

Poellmann, K., Mitterer, H., & McQueen, J. M. (submitted). Use what you can: Storage, abstraction processes and perceptual adjustments help listeners recognize reduced forms. *Frontiers in Psychology*.

Abstract

Three eye-tracking experiments tested whether native listeners recognized reduced Dutch words better after having heard the same reduced words, or different reduced words of the same reduction type and whether familiarization with one reduction type helps listeners to deal with another reduction type. In the exposure phase, a segmental reduction group was exposed to /b/-reductions (e.g., *minderij* instead of *binderij,* 'book binder') and a syllabic reduction group was exposed to full-vowel deletions (e.g., *p'raat* instead of *paraat,* 'ready'), while a control group did not hear any reductions. In the test phase, all three groups heard the same speaker producing reduced-/b/ and deleted-vowel words that were either repeated (Experiments 1 & 2) or new (Experiment 3), but that now appeared as targets in semantically neutral sentences. Clear word-specific learning effects were found for vowel-deletions, but not for /b/-reductions. Generalization of learning to new words of the same reduction type occurred only if the exposure words showed a phonologically consistent reduction pattern (/b/-reductions). In contrast, generalization of learning to words of another reduction type occurred only if the exposure words showed a phonologically inconsistent reduction pattern (the vowel deletions; learning about them generalized to recognition of the /b/-reductions). In order to deal with reductions, listeners thus use

various means. They store reduced variants (e.g., for the inconsistent vowel-deleted words) and they abstract over incoming information to build up and apply mapping rules (e.g., for the consistent /b/-reductions). Experience with inconsistent pronunciations leads to greater perceptual flexibility in dealing with other forms of reduction uttered by the same speaker than experience with consistent pronunciations.

Introduction

In casual speech, speakers tend to articulate in a sloppy way. They frequently reduce words by slurring and even omitting segments or syllables (Ernestus, 2000; Johnson, 2004; Mitterer & McQueen, 2009; Patterson, LoCasto, & Connine, 2003). A given native Dutch speaker may for example reduce the /b/ in *bandiet* 'bandit' to [m] or leave out the first vowel in *kanaal* 'canal' (Schuppler, Ernestus, Scharenborg, & Boves, 2011). Listeners might get used to such pronunciation habits; they may recognize a reduced word better the second time and they may be able to adjust rapidly to new forms of reduction produced by the same speaker. The present study investigates whether listeners adapt to a given reduction type (/b/-reductions or full-vowel-deletions) and whether they can apply their knowledge to previously unheard reduced words of the same and of a different reduction type. Put another way, the present study tests word-specific learning effects as well as generalization of learning within and across reduction types.

Listeners are usually not aware that they encounter numerous reduced word forms every day (Ernestus & Warner, 2011; Kemps, Ernestus, Schreuder, & Baayen, 2004). They use the information provided by the sentence context or also the wider discourse context to predict and, if necessary, restore the upcoming word (Brouwer, Mitterer, & Huettig, 2013; Ernestus, Baayen, & Schreuder, 2002). On a lower level, listeners are also able to exploit the fine phonetic detail present in reduced forms to distinguish for instance between a reduced form [spɔːt] of *support* and the unreduced form [spɔːt] *sport* (Manuel, 1992). Another mechanism which listeners may use to recognize reduced forms better is adaptation, as perceptual learning may be especially important when the conditions for spoken-word recognition become challenging.

Adaptation, for instance, has been found to play a crucial role in recognizing regional and foreign-accented speech (Clarke & Garrett, 2004; Floccia, Goslin, Girard, & Konopczynski, 2006; Mitterer & McQueen, 2009). Listeners are able to adapt rapidly to these deviant pronunciations and can apply their acquired knowledge to novel words (Witteman, Weber, & McQueen, 2013).

The present study tests whether a similar adaptation process also takes place when listeners encounter reduced words in their native language. Like regional and foreign-accented words, reduced words are also variants of canonical pronunciations, but the reduction types chosen for investigation in the present study (/b/-reductions and full-vowel-deletions) were not regionally marked. In contrast to regional and foreign accents, reductions affect predominantly unstressed segments and syllables. They are therefore probably less salient. This might make it harder for listeners to adapt to reduced speech than to regional or foreign-accented speech.

The present study investigates potential adaptation processes and their possible constraints. Consider a Dutch listener hearing the word *paraat* ‘ready’ pronounced as *p’raat*. Different patterns of adaptation are possible that vary in how general they are. First, no adaptation whatsoever may be found. Second, the listener may find it easier to recognize a second instance of the same word with the same reduction pattern. This would be similar to the recognition benefits for words repeated in the same voice that provide some of the evidence for episodic models of word recognition (Goldinger, 1996, 1998; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994). Third, listeners may learn that this speaker deletes vowels in unstressed syllables. This abstractionist learning may be quite specific, so that only very similar reductions benefit (e.g., *Parijs* ‘Paris’ produced as *P’rijs*; note that the Dutch rendition is stressed on the second syllable) or it may include reductions of unstressed vowels in other contexts (e.g., *kanaal* ‘canal’ produced as *k’naal*). The strongest possible generalization may be that the listener assumes that this speaker reduces a great deal, and hence then finds it easier to recognize any kind of reduction uttered by the speaker.

Finding a word-specific learning effect, that is, better recognition of a reduced word on hearing it for the second time compared to the first time, would be evidence for episodic storage of reduced forms. In contrast, observing generalization of learning to new words of the same reduction type (e.g., generalization from *p'raat* to *P'rijs* or *k'naal*) would indicate that an abstraction process is taking place, and that it occurs at a prelexical level. Storing reduced forms alone cannot account for easier recognition of previously unheard reduced words (Cutler, Eisner, McQueen, & Norris, 2010; McQueen, Cutler, & Norris, 2006). In a purely episodic account of lexical access, there is no way to adjust weights of sublexical units like segments and syllables to build up rules that capture regular reduction processes (e.g., *"Potentially restore a bilabial nasal in an unstressed syllable to a bilabial voiced stop if followed closely by another nasal"*). Finding generalization of learning to new words of the same reduction type would thus support the claim that there is abstraction in lexical access. Observing generalization of learning from one reduction type to another may also be evidence for abstraction—if there is enough similarity between the reduction types to abstract over the respective mapping rules. Consider, for example, two types of prefix reductions, such as *ge*- /ɡə/ → /ɡ/ and *be*- /bə/ → /b/ in German. An abstraction rule may be: "*Potentially insert a schwa after an initial voiced stop"* (instead of "... *after an initial voiced velar/bilabial stop"*). However, should generalization of learning across reduction types be found for very different reduction types, such as the /b/-reductions and full-vowel-deletions examined here, this would more likely indicate a non-specific adjustment and be evidence for the flexibility of the perceptual system. That is, instead of specific adaptation processes (storage of reduced forms and/or abstraction of reduction rules), listeners could make a more general adjustment to the current talker's speaking style.

To test these possible adaptation effects, the printed-word eye-tracking paradigm (McQueen & Viebahn, 2007) was used. In the exposure phase, one group of participants was exposed to segmental reductions, another group was exposed to syllabic reductions and a third group was exposed only to canonical pronunciations. The first group, the segmental reduction group, heard /b/-reductions, where the word-initial /b/ was reduced to a bilabial nasal (e.g., *minderij* instead of *binderij* 'book binder'). The second group, the syllabic reduction group, heard words in which the first, unstressed full vowel was deleted (e.g., *p'raat* instead of *paraat* 'ready'). The third group, the control group, heard the same words but all in unreduced form.

These two reduction types were chosen to examine adaptation to two different-sized linguistic units, the phoneme and the syllable, and the possible interaction of the adaptation effects. An earlier study showed that listeners adapt to syllabic reductions involving a morpheme: After exposure to words containing the reduced prefix *ver*- (realized as [fː]), Dutch listeners recognized previously unheard reduced *ver*-words better than a control group (Chapter 3). In the present study, we test whether this is also the case for non-morphemic syllables. The deletion of the unstressed, full vowel in CVC-initial words like *paraat* always led to a reduction in the number of syllables, which is why this reduction type was called 'syllabic'. A pure comparison of morphemic and non-morphemic reductions, however, turned out to be impossible in Dutch. Ideally, one would like to compare a morphemic reduction type (that only affects one specific morpheme, i.e., the same strings of segments, such as Dutch *ge*-) to a non-morphemic reduction type that also only affects one specific string of segments (e.g., *pa*-). The Dutch lexicon, however, does not contain enough words starting with one specific unstressed non-morphemic syllable to conduct such an experiment. This constraint on the (non-)morphemic status hence leads inevitably to higher variability in the segmental structure of the vowel-targets compared to the *ver*-targets examined in Chapter 3. The degree of consistency with which words are reduced may determine which adaptation processes (e.g., storage, abstraction rules, general flexibility) listeners are able to use.

In the test phase, all three groups heard /b/-reductions and vowel-deletions. The reduced words were either the same as in the exposure phase (in Experiments 1 and 2) or different (in Experiment 3). If listeners adapt to a given reduction type and if they can transfer this knowledge to new words (Experiment 3) and/or to other reduction types (Experiment 1-3), participants in the experimental groups should recognize reduced words better than participants in the control group.

Regardless of specifics concerning the reduction (such as size of the reduced unit or input consistency), it seems plausible that a reduced word can be recognized more easily if it is encountered a second time. We therefore expect to find word-specific learning effects for both /b/-reductions and vowel-deletions.

Moreover, we predict that learning about /b/-reductions generalizes to new words that are reduced in the same way. Such generalization effects have been observed for a similar kind of /b/-reduction where the word-initial voiced stop was reduced to a labio-dental approximant [ʋ] (Chapter 3) and for learning about segmental idiosyncrasies (McQueen, et al., 2006). In the McQueen, et al. (2006) study, listeners adapted to an ambiguous sound (between /s/ and /f/) and transferred their knowledge to previously unheard minimal pairs that only differed in containing either /s/ or /f/.

The predictions concerning within-reduction-type generalizations for full-vowel-deletions are less clear. The constraint on the (non-)morphemic status of the syllable leads to higher variability in the segmental structure of the vowel-targets compared to the /b/-targets. If the input has to be highly consistent for the creation of abstract mapping rules, we might not observe generalization of learning.

The two reduction types under investigation differ in several respects, such as the degree of reduction (weakening of the [b] vs. deletion of the vowel), in the segment that is reduced (bilabial voiced stop vs. full vowel) and in the position the reduced segment occurs (first position for /b/-reductions vs. second position for vowel-deletions). In order to observe generalization of learning across reduction types, listeners would hence have to adapt on a fairly global level. However, such global adjustments to challenging listening conditions have been observed before (Brouwer, Mitterer, & Huettig, 2012; McQueen & Huettig, 2012).

In order to assess the frequency with which our chosen reduction types occur in spontaneous speech, we conducted a corpus study following the principles of Pluymaekers, Ernestus, and Baayen (2005). First, all sound files containing a /b/-initial word with a nasal in third position and an unstressed first syllable were extracted from the Corpus of Spoken Dutch (Oostdijk, 2000). Per word type (this notion here not only describes words belonging to different lemmas but also different word forms of one lemma, e.g., an inflected verb form or the plural of a noun) only one token was randomly chosen to determine its phonetic realization. Out of 65 word types, six showed a /b/ → [m] reduction in the first segment (i.e., 9.2% of the considered cases). A similar analysis was conducted to assess the frequency of full-vowel-deletions in initially unstressed words. The vowel was deleted in eight out of 66 word types (i.e., in 12.1%) containing either a voiceless plosive (/p/, /k/) or a voiceless velar fricative (/x/) in first position and an alveolar nasal or liquid in third position. This was also the segmental structure used in Experiments 1-3 for the CVC-initial words. The chosen reduction types were thus indeed a real-world phenomenon and comparable in terms of frequency.

Experiment 1

The aim of Experiment 1 was to test whether listeners are able to recognize segmental and syllabic reductions better, if they already encountered the same words in reduced form before. Experiment 1 also asked whether learning about reductions might generalize from one reduction type to another (i.e., from /b/-reductions to full-vowel deletions and/or vice versa). In the exposure phase, one group was exposed to /b/-reductions (segmental reduction group), a second group was exposed to full-vowel deletions (syllabic reduction group), while a third group was exposed to canonical forms only (control group). In the test phase, all three groups were tested on reduced-/b/ words and vowel-deleted words. Importantly, these reduced words had already occurred in reduced or canonical form (depending on the group) in the exposure phase. If listeners can adapt to reduced words, the segmental reduction group should recognize the reduced-/b/ words better than the syllabic reduction group and the control group because of their previous exposure to these words in reduced form. The same holds for participants in the syllabic reduction group: If they can adapt to vowel-deleted words, they should perform better on these words than the segmental reduction group and the control group. If listeners can additionally transfer their knowledge about one reduction type to another, the segmental reduction group should outperform the control group on the vowel-deleted words and the syllabic reduction group should outperform the control group on the reduced-/b/ words.

Method

Participants. Seventy-five participants of the Max Planck Institute's subject pool, all native speakers of Dutch, were paid to take part. All reported normal hearing and normal or corrected-to-normal vision.

Design. Participants were randomly assigned to one of three groups: a segmental reduction group, a syllabic reduction group and a control group. They listened to sentences, saw four printed words on a computer screen and were asked to click on the word that occurred in the sentence. In the exposure phase, these words were reduced or not, depending on the participant's group (see Table 1). Participants did not see these potentially reduced words on the screen. Instead, they saw (and had to click on) words that occurred later in the sentences. All three groups were also exposed to unreduced /m/- and unreduced consonant-cluster-words (e.g., /mɑtros/ *matroos* 'sailor' and /knɔflok/ *knoflook* 'garlic').

In the test phase, all three groups heard reduced /b/-words and vowel-deleted words in the experimental trials. These were the same words as had appeared in the exposure phase (e.g., [mɪndərɛɪ] instead of [bɪndərɛɪ] *binderij* 'book binder' and [prat] instead of [parat] *paraat* 'ready'). All groups also heard new canonical /m/- and new canonical consonant-cluster words in the unreduced condition. The reduced /b/-words, the vowel-deleted words, the unreduced /m/- words and the consonant-cluster words were the targets and were therefore displayed on the computer screen in (canonical) orthographic form.

Table 1. Experimental design and types of stimuli in Experiment 1 and 2. The potentially reduced /b/-initial words and the vowel-deleted words of the exposure phase were repeated in reduced form in the test phase.

	Condition	Canonical word-form	Segmental reduction group	Syllabic reduction group	Control group
			/b/ → [m]	full V deletion	no reduction
Exposure phase	Experimental	/bɪndərɛɪ/	[**m**ɪndərɛɪ]	[bɪndərɛɪ]	[bɪndərɛɪ]
	Unreduced	/mɑtros/	[mɑtros]	[mɑtros]	[mɑtros]
	Experimental	/parat/	[parat]	[**pr**at]	[parat]
	Unreduced	/knɔflok/	[knɔflok]	[knɔflok]	[knɔflok]
Test phase	Experimental	/bɪndərɛɪ/		[**m**ɪndərɛɪ]	
	Unreduced	/murɑs/		[murɑs]	
	Experimental	/parat/		[**pr**at]	
	Unreduced	/xlɑns/		[xlɑns]	

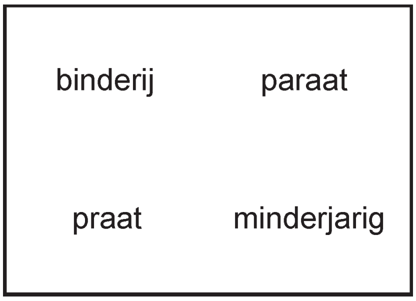
Materials. The target words (i.e., the words participants had to click on) appeared towards the end of spoken sentences. Each target word occurred in a different sentence context not containing any further /b/s in unstressed syllables or any further unstressed CVC-sequences which would result in legal consonant clusters when omitting the vowel. In the exposure phase, the sentence context predicted the target word. The potentially reduced item, however, occurred before the target word in the experimental trials (e.g., *Pas in een <u>[b]/[m]inderij</u> wordt een boek of* ***tijdschrift*** *afgemaakt* 'Only at a <u>book binder</u>, a book or **magazine** gets finished', where bold font indicates the target word and underlining marks the potentially reduced critical item). This was done to prevent participants from clicking on the same words twice, once in the exposure phase and once in the test phase. In the test phase, the semantic contexts preceding the target words were kept uninformative (e.g., *Het tekstverwerkingprogramma kende het woordje* ***<u>[m]inderij</u>*** *niet* 'The word processor did not know the word **<u>book binder</u>**'). During each sentence, there were always four printed words on the screen. In the test trials, these were a /b/-word, a /m/-word, a vowel-intact word and a consonant-cluster word (see Figure 1 for an example display).

For the test phase, 24 target-competitor pairs for each type of target word (/b/-target and vowel-target in the experimental condition and /m/-target and consonant-cluster-target in the unreduced condition) were selected (see Table A1 for the /b/-targets and the vowel-targets, and their respective competitors). If a /b/-word was the target, a /m/-word was the competitor and vice versa. The same holds for vowel- and consonant-cluster-targets. All /b/- and /m/-initial words contained an unstressed first syllable. In second position, any vowel including schwa could occur followed by a nasal in third position. The latter condition was necessary for all /b/-targets to motivate nasalization at the beginning of the word. However, there are not sufficient /m/-initial words in Dutch containing a nasal in third position to create perfectly matched pairs of /b/-targets and /m/-competitors. Ideally, /b/-words and /m/-words should be as similar as possible with as much overlap in the reduced forms as possible (e.g., *binderij* 'book binder' pronounced as [mɪndərɛɪ] overlaps in the first two syllables with [mɪndərjarəx] *minderjarig* 'underage'). Due to the infrequent occurrence of a nasal in third position following an /m/ in first position, the /m/-targets contained a random consonant in third position (and so did the corresponding /b/-competitors; e.g., *moeras* 'swamp' and *boerin* 'farmer's wife'). Target-competitor pairs were further matched in terms of number of syllables, stress pattern and word frequency as much as possible (see Table A1).

The principles of as much overlap and similarity between targets and competitors as possible also applied to the (reduced) vowel- and (unreduced) consonant-cluster-words. Vowel-words started with an open syllable, consisting of a voiceless consonant (either /p/, /k/, or /x/) and a full vowel, followed by a liquid or /n/ in third position (e.g., *paraat* 'ready'), so that the sequence resulting from vowel deletion would be phonotactically legal in Dutch. The consonant-cluster words started with the same voiceless consonants directly followed by a liquid or [n] (e.g., *praat* 'talk'). While the stress of the vowel-words was on the second syllable, the consonant-cluster-words were stressed on the first syllable, so that both word types were matched on stress pattern when the full vowel of the vowel words was deleted (e.g., p'RAAT for paRAAT 'ready' and PRAAT 'talk'). Again, target-competitor pairs were matched on number of syllables (in the reduced form) and word frequency (see Table

A1). The only constraint for the 24 target-"competitor" pairs per word type in the exposure phase was that they did not overlap.

Figure 1. Example display of a test trial in Experiment 1, 2, and 3.



Stimulus construction. Digital recordings of the stimuli were made by a female native speaker of Dutch in a sound-proof booth, sampling at 44.1 kHz. She was instructed to produce the sentences in a casual way, not just reading them aloud. For sentences containing canonically pronounced /b/-targets, an additional set containing reduced forms was created by replacing the /b/ with an /m/ from a word with the same vowel context. The spliced parts were adjusted in pitch (with PSOLA in PRAAT, Boersma & Weenink, 2010) and intensity to their new context. The transitions in amplitude preceding and following the spliced-in [m]s were smoothed where necessary in order to reduce splicing artifacts. The set of sentences containing reduced-vowel words was created by cutting out the first (unstressed) vowel of the recorded versions of these words with intact vowels. Sentence contexts were thus identical across the reduced and unreduced forms of each target word. Sentences containing /m/- and consonant-cluster-targets were not manipulated.

Procedure. Participants were seated in a sound-attenuated booth at a comfortable viewing distance from the computer screen. Eye movements were monitored using an SR Research EyeLink 1000 set-up, sampling at 1 kHz. The auditory stimuli were presented to the participants over headphones. Prior to the experiment, participants received written instructions that informed them that they would see four printed words on the screen and asked them to click on the word that occurred in the sentence.

At the beginning of each trial, a fixation cross appeared in the centre of the screen for 500 ms. Four printed words (in a 25-point Arial font) were then presented. After 1500 ms, the auditory stimulus was played. As soon as participants had listened to the entire sentence and had clicked with the mouse on the screen, the following trial was initiated. Every ten trials, a drift correction was carried out. Participants had the opportunity to take a break after every 50th stimulus. The experiment started with six practice trials. The randomized order of the 96 exposure and 96 test trials was different for each participant. An experimental session took approximately 25 minutes.

Results

Exclusion criteria. Mouse click responses (reaction time and accuracy data) and eye movements served as dependent variables. For the eye-tracking data, we analyzed the data from the participant's right eye. For the analysis of the eye-tracking data, a total of 2.9% of the trials were excluded, because participants either appeared to have looked away from the screen (2.0%) or failed to click on the target or the potentially confusable competitor (0.9%). Clicks on the competitor were not excluded from all of the analyses, as the competitors sometimes better fitted the exact auditory input with reduced forms than the targets. For instance, reduced *p'raat* better fitted the canonical form of the competitor *praat* than the canonical form of the target *paraat*. Furthermore, the semantics of the test sentences did not make clear which word was the target. In the case of minimal pairs such as *paraat* and *praat*, participants thus never received disambiguating information about which of the two words they should click on. Therefore, clicks on competitors were not regarded as errors in the analyses of the eye-tracking and the reaction time data. Note also that excluding trials from the eye-tracking analysis in which participants clicked on the competitor would invalidate any learning effects. Presumably, participants look more at the competitor when they click on it. Excluding these trials would result in a greater preference for the target over the competitor and would thus misleadingly indicate a greater learning effect than was actually present. Moreover, the focus in the RT analyses is on the comparisons across the three exposure groups; these comparisons are thus orthogonal to any differences between targets and competitors. Click responses to competitors, however, were regarded as incorrect in the analysis of the accuracy scores.

Table 2 displays descriptive statistics on RTs for trials in which participants clicked either on the target or on the phonological competitor in the test phase. Participants in the syllabic reduction group took longer to respond than participants in the segmental or no-reduction group. Participants, however, were not asked to respond as fast as possible. Some participants chose to do so; others waited for the sentence to finish before giving a response. The high standard deviation (sd) values reflect these different strategies. Extreme cases, that is, trials in which participants responded either too fast or too slowly, were also excluded. To do that, a linear mixed-effects model containing only participants and items as random effects and Trial Number as fixed effect was run. The residuals of this atheoretical model were computed. Based on visual inspection of a residual plot, 19 trials (0.5%) in the test phase (with residuals either below -1300 ms or above 3200 ms) were excluded.

Table 2. RTs in ms in the test phase of Experiment 1 for clicks on targets and competitors.

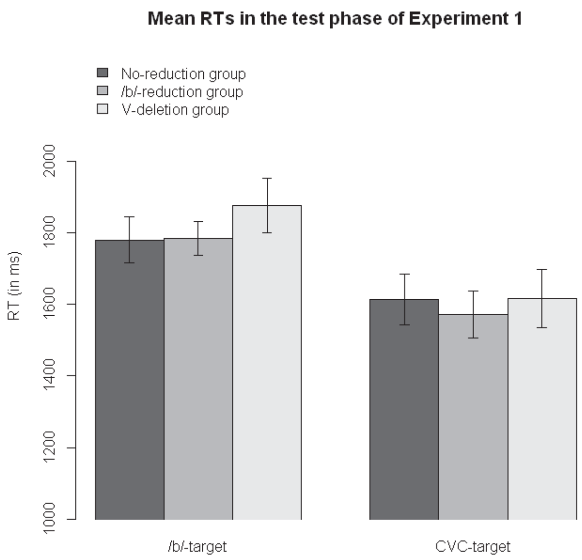
RT in ms	Segmental reduction group	Syllabic reduction group	No reduction group
mean	1695	1769	1714
sd	687	803	751
min	455	294	555
max	7303	8647	9107

Statistical testing. Linear mixed-effects models were used to analyze the click responses and the eye movement data on the experimental trials (containing /b/-targets and CVC-targets). In the latter case, the proportion data was transformed using the empirical logit function. Participants and Items were entered in the model as random factors including random slopes for Items. Group served as fixed effect. The segmental reduction condition (/b/-words) and the syllabic reduction condition (vowel-words) were analyzed independently. Trial Number was entered as another fixed effect with values centered around zero in the models for the accuracy and RT data. This variable was added to account for additional variance, as task performance often improves over the course of an experiment. The results for Trial Number, however, will not be reported below. Thus, we tested whether RTs, accuracy scores and target preference (as determined by the difference between proportion of target and competitor fixations) for the reduced words were influenced by the fixed effect of

Group. That is, we examine whether the groups differ in how fast and how accurate they recognize the reduced /b/-words and the vowel-deleted words and whether they show different target-competitor preferences when they process reduced words. The control group was always mapped on the intercept, so that the analysis gives two regression weights for the factor Group, one for the difference between the control group and the segmental reduction group and one for the difference between the control group and the syllabic reduction group. For the eye-tracking analyses, we used sliding 200 ms time windows from 200-1500 ms after target onset starting at every 100 ms.

Test Phase. ***Reaction Time data.*** Figure 2 displays the mean RTs of all three groups for the reduced /b/-words and the vowel-deleted words in the test phase of Experiment 1. In the segmental reduction condition (/b/-targets), all three groups responded about equally fast and no significant differences between the groups emerged. In the syllabic reduction condition (CVC-targets), there was also no main effect of Group. That is, neither of the experimental groups responded faster than the control group to the reduced words. We thus did not observe any adaptation effect in the RT data.

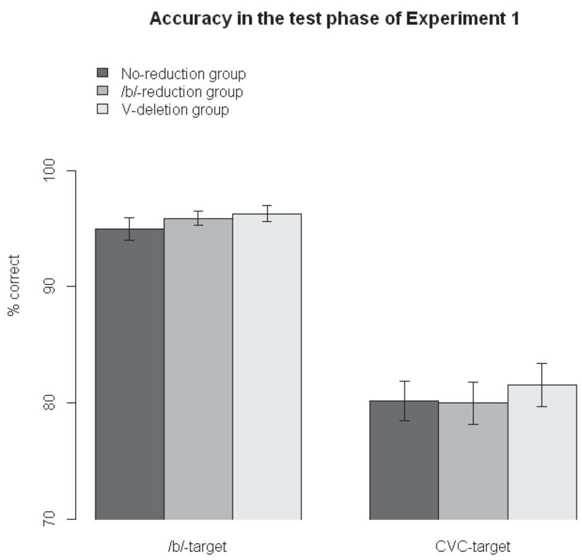
Figure 2. Mean RTs and SEs in the test phase of Experiment 1.



Accuracy data. The accuracy data in the test phase of Experiment 1 are displayed in Figure 3 in terms of percentage correct click responses and SEs. In the segmental reduction condition, the main effect of Group was significant. Both the segmental reduction group ($b_{\text{Segmental reduction group}}$ = 3.4, p < .001) and the syllabic reduction group ($b_{\text{Syllabic reduction group}}$ = 2.3, p < .001) gave more correct responses to /b/-targets than the control group. We thus observed an adaptation effect for both experimental groups in the accuracy data of the segmental reduction condition.

In the syllabic reduction condition (CVC-targets), the main effect of Group was not significant. That is, neither of the experimental groups differed from the control group. We thus did not observe a significant adaptation effect for either group.

Figure 3. Accuracy in % correct click responses and SEs in the test phase of Experiment 1.



Eye movement data. The eye movement patterns for the segmental reduction condition (/b/-targets) of the two experimental groups compared to the no-reduction control group are displayed in Figure 4. Early on, in a descriptive time window from 200ms - 500ms after target onset, the control group (represented by black lines) looks more often to the competitors (dashed lines) when hearing a reduced /b/-word than the segmental reduction group (in grey, Panel A) or the syllabic reduction group (in grey, Panel B). From around 500 ms onwards, all three groups show a similar preference for the /b/-targets (solid lines).

Statistical analyses showed that the difference in target-competitor preference between the segmental reduction group and the control group did not reach significance. The main effect of Group, however, was marginally significant for the syllabic reduction group in the time window from 300-500 ms after target onset ($b_{\text{Syllabic reduction group}} = 1.9$, $p = .06$). That is, we observed a weak adaptation effect for the syllabic reduction group in the segmental reduction condition, hence a weak generalization of learning across reduction types.

Figure 4. Proportion of fixations in the segmental reduction condition of Experiment 1.

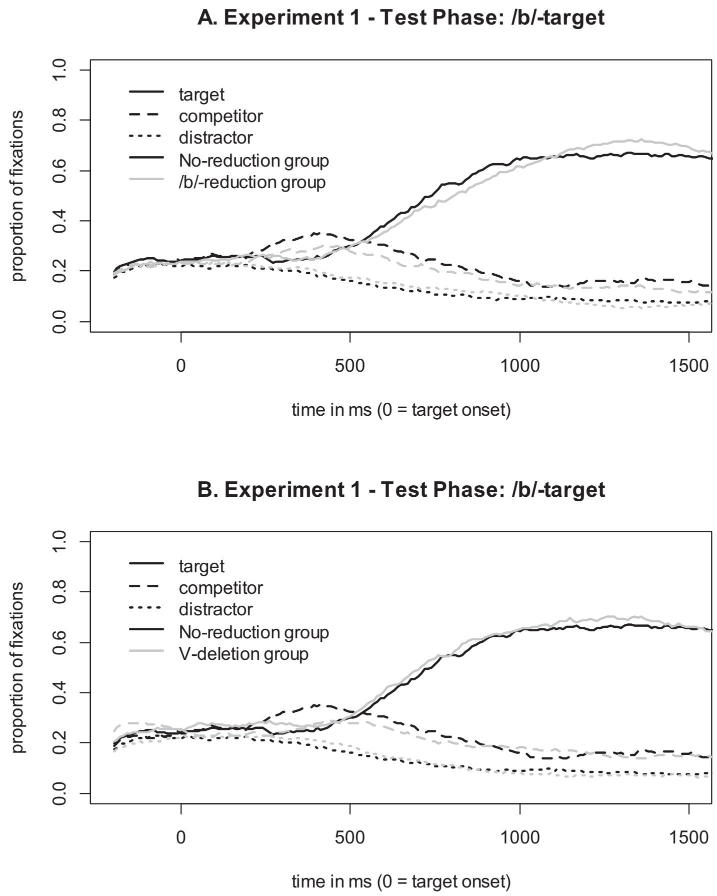
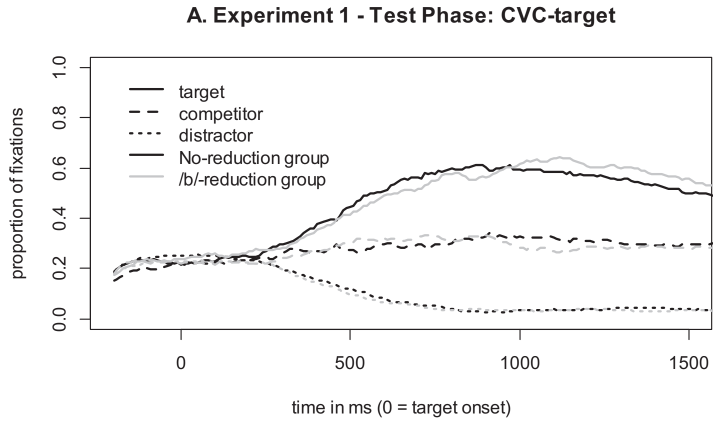
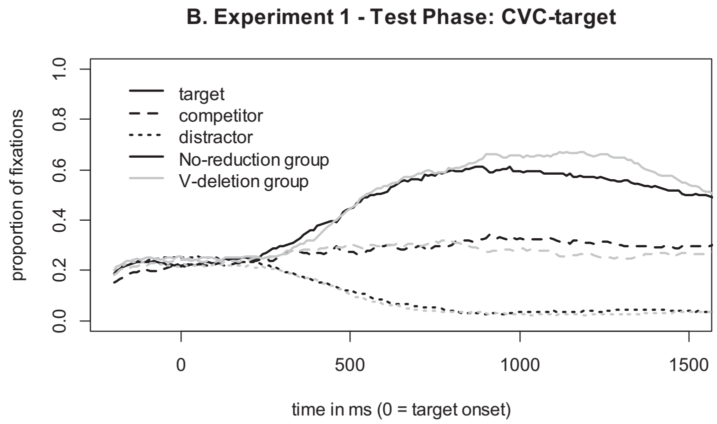


Figure 5 displays the corresponding eye movement data for the syllabic reduction condition (CVC-targets). In the first 900 ms after target onset, all three groups show a very similar pattern for the vowel-deleted words. Only later, the two experimental groups have descriptively a greater target preference for the vowel words than the control group.

Statistical analyses did not reveal a significant difference between the control group and the segmental reduction group, but revealed that the main effect of Group was significant in the time window from 1100 ms - 1400 ms for the syllabic reduction group ($b_{\text{Syllabic reduction group}}$ = 0.93, t = 2.2, p < .05). In this time window, the syllabic reduction group had a greater target-competitor preference for the CVC-words than the control group. For the syllabic reduction group, we found thus an adaptation effect.

Figure 5. Proportion of fixations in the syllabic reduction condition of Experiment 1.





Discussion

In Experiment 1, we found adaptation effects for both the segmental and the syllabic reductions. Learning about segmental reductions became evident in the accuracy data, but not in the eye-tracking data. For the syllabic reductions, this pattern was reversed: A learning effect was found in the eye-tracking data, but not in the accuracy data. Moreover, there was also evidence of generalization of learning across reduction types. Generalization across reduction types, however, was only found in one direction: learning about vowel deletions generalized to /b/-reductions, as shown by the accuracy data and the eye movement data for the segmental reductions. In contrast, learning about /b/-reductions did not generalize. That is, the segmental reduction group could not apply their experience with reductions to the vowel-deleted words.

The learning effects found in Experiment 1 seem somewhat weak overall. An explanation for that may be the low predictability of the potentially reduced words in the exposure phase. Participants did not see the potentially reduced words on the computer screen and these words appeared early in the sentences. Participants may therefore not have been able to predict them. Having information about the upcoming reduced words in advance could however facilitate learning. Experiment 2 was run to test this hypothesis.

Experiment 2

Experiment 2 tested whether providing lexical information about the reduced words in the exposure phase might strengthen the learning effects found in Experiment 1. Therefore, we changed the exposure sentences for the reduced /b/-words and the vowel-deleted words, leaving the exposure sentences for the /m/-words and the consonant-cluster words intact. The sentence contexts now predicted the potentially reduced words. To avoid the orthographic versions of the reduced words appearing twice on the screen, eye-tracking was not used in the exposure phase. Instead, participants simply listened to the exposure sentences and were asked to answer questions about the content of some of the sentences (those in the unreduced condition containing /m/- or CC-words).

The test phase was kept the same as in Experiment 1, apart from minor changes in three sentences (see Methods). Further purposes of Experiment 2 were to replicate the generalization effect from vowel-deleted words to reduced /b/-words found in Experiment 1 and to test whether, with predictable sentences, a generalization effect in the other direction (from reduced /b/-words to vowel deletions) might also occur.

Method

Participants. Sixty Dutch participants of the Max Planck Institute's subject pool, none of whom had participated in Experiment 1, were paid for their participation. All had normal hearing and normal or corrected-to-normal vision.

Design. The design was similar to that in Experiment 1. The main difference was a change in task during the exposure phase, where participants had to answer questions regarding the content of some of the reduction-free sentences without their eye movements being tracked.

Materials. While for the unreduced condition (the /m/-words and the CC-words) the same exposure sentences as in Experiment 1 were used, new exposure sentences were generated for the experimental condition (the potentially reduced /b/-words and vowel-deleted words). The critical words now appeared towards the end of the sentences (e.g., *Als een manuscript gedrukt is, moet het naar de [b]/[m]inderij*. 'When a manuscript is printed, it has to go to a book binder') and were predicted by the semantic context (see cloze test below). The materials for the test phase were taken from Experiment 1. Only three target words were changed slightly (*bankier* 'banker' → *bankiers* 'bankers', *benauwen* 'to oppress' → *benauwd* 'sultry', *coulisse* 'wing [of theatre stage]' → *coulissen* 'wings, pl.') so that it was possible to create more natural sentences for the exposure phase.

Cloze tests. Cloze tests were run to check the degree of predictability of the potentially reduced words in the exposure sentences. The 48 sentences were presented in a randomized order with the critical word replaced by a gap. Participants were instructed to complete these sentences with one word. They were asked to type in at least one answer, but had the possibility to give up to seven. After typing in their answer(s), participants saw the same sentence again completed with the corresponding /b/- or CVC- target. They were asked to rate how well the proposed solution completed the sentence context on a scale from 1 ("Word does not fit at all") to 7 ("Word fits perfectly"). The cloze tests were self-paced; it took participants 15 to 30 minutes.

An initial test with eighteen Dutch native speakers of the Max Planck Institute's subject pool, who had not participated in Experiment 1, showed that for some sentences the target word was mentioned in less than 25% of cases. These were improved if possible. A second version of the cloze test was run with 19 new Dutch participants. We analyzed the percentages of mentioned target words in the sentence completion task and the mean ratings for the targets in the rating task. The critical /b/-words were mentioned in 36% of the cases, while the critical vowel-words were mentioned in 51% of the cases. This difference does not reflect a frequency effect, as the /b/-targets are more frequent than the CVC-targets (see Table A1 and A2). But it can possibly be explained by the higher constraints on the initial selection of the /b/-words. Only /b/-words were chosen which had a nasal in third position and for which a /m/-initial competitor with as much onset overlap as possible existed. Similar constraints on the vowel words were less strong, as the consonants in first and third position could vary. Although participants did not come up with our solutions in many cases, they rated those solutions very highly on average: On a scale from 1 to 7, with higher ratings meaning better fits, participants rated the /b/-targets with 6.1 and the vowel-targets with 6.3 on average.

Stimulus construction. The new exposure sentences were recorded by the same female Dutch speaker who provided the stimuli for Experiment 1. The reduced stimuli were created in the same way as described in Experiment 1.

Procedure. Participants were tested in a sound-proof booth. They were told that the experiment consisted of two parts. For the first part, they were asked to listen to sentences that were presented over headphones and to answer questions regarding the content of these sentences (by clicking on one out of two suggested solutions) that might appear at random points in time on the screen.

Each exposure sentence was preceded by 500 ms of silence and followed by 2000 ms of silence. If a question and two possible solutions were to appear on the screen (after six /m/-word sentences and after six CC-word sentences, i.e., in $^{1}/_{8}{}^{th}$ of the exposure trials), they followed the auditory stimulus immediately. After participants had clicked on the screen, it took 1000 ms before the next exposure trial started. The order in which the exposure sentences were played was randomized for each participant individually. Participants had the opportunity to take a break approximately halfway through the experiment, after the 50th stimulus (out of 96).

The procedure of the test phase was identical to the one in Experiment 1, except that eye movements were monitored using an SR Research EyeLink II, sampling at 500 Hz. An experimental session took approximately 30 minutes.

Results

Exclusion criteria. The same criteria as in Experiment 1 were applied for trial exclusion leading to the exclusion of 2.2% of the data due to fixations outside of the screen area and of another 1.3% due to failure to click on the target or the potentially confusable competitor. An additional 0.5% of trials were discarded because they were considered to be RT outliers (with residual values either below -2300 ms or above 3100 ms). For the eye-tracking data, we analyzed the data from the better eye of the participants (i.e., the eye that showed less error in the validation of the calibration of the eye-tracker).

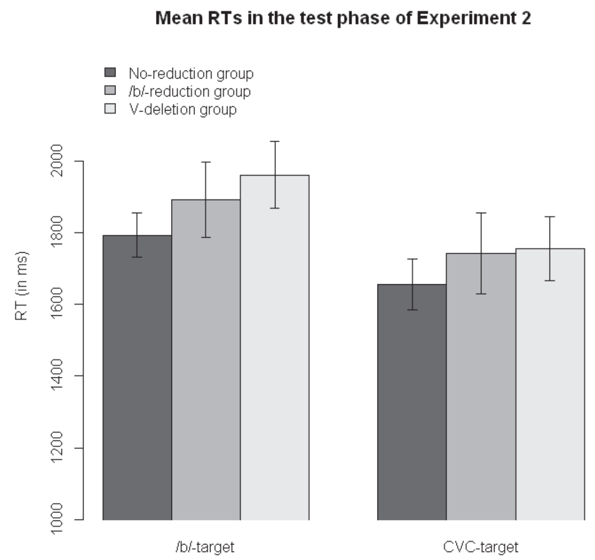
Exposure Phase. Participants of all groups hardly made errors in the comprehension questions of the exposure phase. Each group obtained a score of 99% correct responses.

Test Phase. ***Reaction Time data.*** Table 3 shows the descriptive statistics for the RT data in the test phase of Experiment 2. The mean RTs and their SEs of all three groups for the reduced /b/-words and vowel-deleted words are displayed in Figure 6. The no-reduction control group seems to respond slightly faster than the two experimental groups in both the segmental reduction condition (/b/-targets) and the syllabic reduction condition (CVC-targets). However, the main effect of Group was not significant in either condition. As there was no main effect of Group in the RT data indicating that one or both of the experimental groups responded faster to the reduced targets than the control group, we did not observe any adaptation effect.

Table 3. RTs in ms in the test phase of Experiment 2 for clicks on targets and competitors.

RT in ms	Segmental reduction group	Syllabic reduction group	No reduction group
mean	1898	1880	1730
sd	1597	868	660
min	346	512	455
max	34467	9863	6384

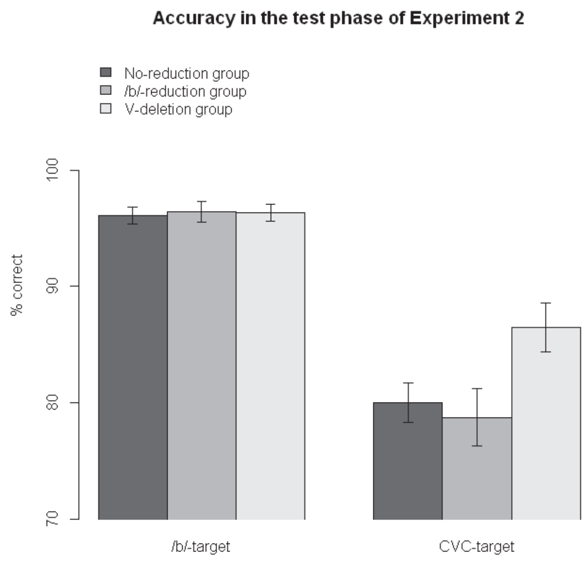
Figure 6. Mean RTs and SEs in the test phase of Experiment 2.



Accuracy data. Figure 7 shows the accuracy data in percentages correct responses and SEs of all three groups for the reduced /b/-words and vowel-deleted words. All three groups performed near ceiling in the segmental reduction condition (/b/-targets). There was no difference between the groups indicating that the experimental groups did not respond more accurately than the control group. We did thus not observe any adaptation effect in the accuracy data for the segmental reduction condition.

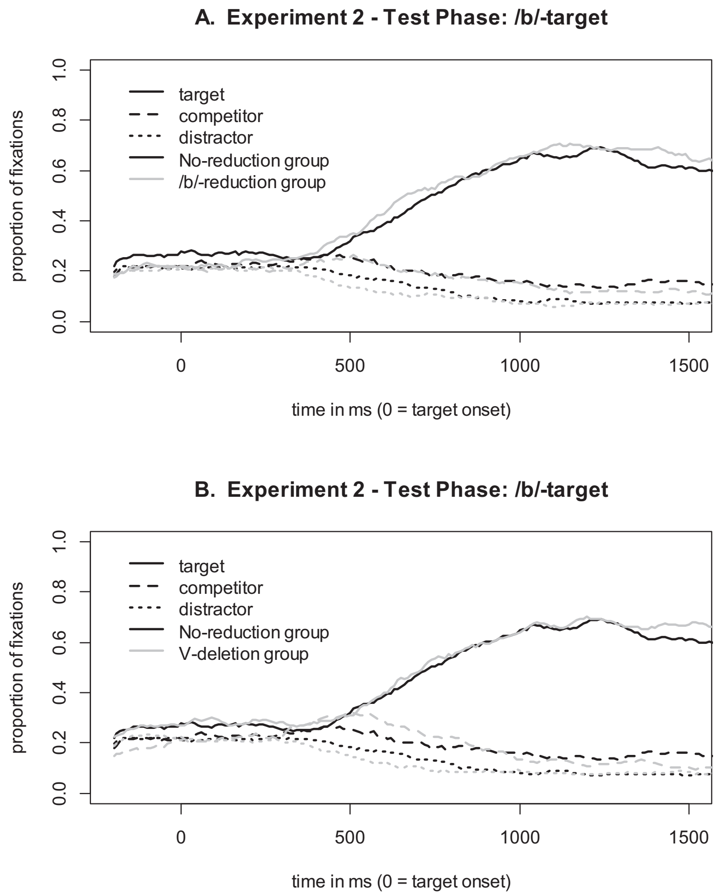
In the syllabic reduction condition (CVC-targets), the main effect of Group was significant for the syllabic reduction group ($b_{\text{Syllabic reduction group}} = 0.9$, $p < .01$), but not for the segmental reduction group ($b_{\text{Segmental reduction group}} = 0.2$, $p > .1$). That is, only the syllabic reduction group gave more correct answers when hearing a vowel-deleted word than the no-reduction control group. We thus observed a learning effect for the syllabic reduction group, but no generalized learning effect for the segmental reduction group.

Figure 7. Accuracy in % correct click responses and SEs in the test phase of Experiment 2.



Eye movement data. Figure 8 shows the eye-movement patterns in the segmental reduction condition for the segmental reduction group (in grey, Panel A) and the syllabic reduction group (in grey, Panel B) compared to the no-reduction control group (in black). All three groups behave very similarly when hearing reduced /b/-words. There was indeed no main effect of Group. That is, we did not observe any learning effect for the segmental reduction condition in the eye-tracking data.

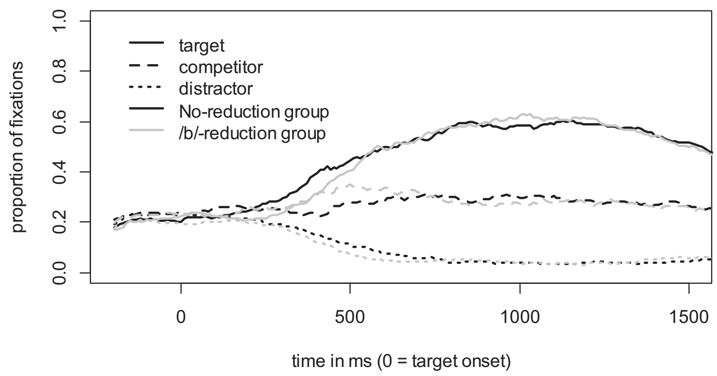
Figure 8. Proportion of fixations in the segmental reduction condition of Experiment 2.



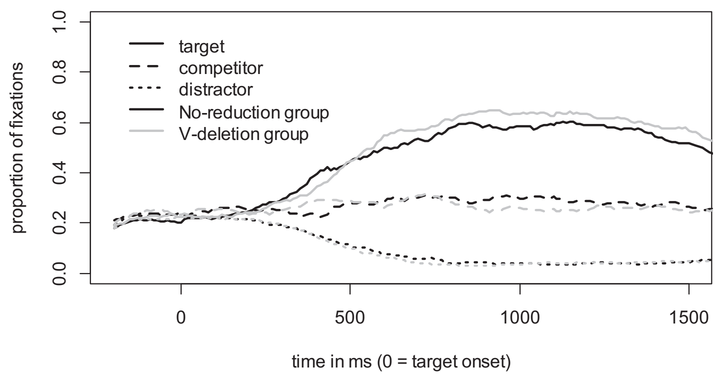
The corresponding eye movement data for the syllabic reduction condition (CVC-targets) are displayed in Figure 9. Statistical analysis revealed a marginal main effect of Group ($b_{\text{Segmental reduction group}}$ = -0.64, t = -2.0, p = .06) in the time window from 200 ms - 500 ms after target onset. The segmental reduction group had a smaller preference for the CVC-targets over the CC-competitors than the control group in this time window. We thus observed a marginal inhibitory effect for the segmental reduction group, given that participants in this group, who had experience with another type of reduction, showed a smaller target preference than participants in the control group, who had not been exposed to any reductions. Furthermore, no learning effect was found for the syllabic reduction group.

Figure 9. Proportion of fixations in the syllabic reduction condition of Experiment 2.

A. Experiment 2 - Test Phase: CVC-target



B. Experiment 2 - Test Phase: CVC-target



Discussion

Experiment 2 was conducted to replicate the findings of Experiment 1 and to test whether predictability of the reduced words during exposure enhances the learning effects. As in Experiment 1, adaptation was observed in the syllabic reduction condition. Contrary to the previous experiment, it was found in the accuracy data, not in the eye-tracking data. The pattern of target and competitor fixations for the syllabic reduction group, however, was in the expected direction (see Figure 9B). We did not replicate the learning effect for segmental reductions found in the accuracy data in Experiment 1. Neither could we replicate the generalized learning effect for the syllabic reduction group for vowel-deletions to /b/-reductions (that was also evident in the accuracy data of Experiment 1). Another "generalization" effect, however, emerged. In contrast to Experiment 1, the segmental reduction group differed from the control group when dealing with vowel deletions. In the eye-tracking data, they showed a smaller target-competitor preference for CVC-targets. That is, even though they did not show a learning effect for /b/-reductions, participants in the segmental reduction group seemed to be hindered by their exposure to /b/-reductions and struggled more with recognizing the vowel-deleted words than the control group.

In Experiments 1 and 2, we found learning effects for repeated /b/-reductions and vowel-deletions. At this point, we cannot say yet whether these effects are truly word-specific, meaning that they arose because the reduced forms were stored after their first encounter in the mental lexicon and then accessed again as they were encountered the second time in the test phase. The observed effects could also have arisen because of rule abstraction. To determine which mechanism is responsible for the learning effects found for repeated reduced words in Experiments 1 and 2, we test whether learning can generalize to other words of the same reduction type in Experiment 3. If there is no or only weak evidence for generalized learning, then the effects found for repeated words are very likely to be word-specific. In contrast, if there is strong evidence for generalized learning, then the effects found for repeated words are likely due to abstraction processes.

The null result for the segmental reductions in Experiment 2 suggests that predictable sentences alone might not be enough to induce a stable adaptation effect. In Experiment 3, we therefore combined aspects of the exposure phase of Experiment 1 (eye-tracking with printed words on the screen) with aspects from Experiment 2 (predictable sentence context). This procedure should render the reduced target words highly predictable, which in turn could lead to a strong learning effect. Using eye-tracking in the exposure phase can tell us whether participants actually make use of the sentence context (i.e., they might already look at the target word before it is mentioned).

Experiment 3

In Experiment 3, we tested whether learning about reductions can generalize across words (within a reduction type). To that end, new /b/-words and new vowel-words were selected for the exposure phase and new exposure sentences were created in which those words were predictable. In the exposure phase, participants had to click on the potentially reduced /b/- and vowel-words, while their eye-movements were recorded. The test phase was the same as in Experiment 1. Importantly, the target words of the test phase did not occur in the exposure phase. Apart from the generalization of learning within a reduction type, Experiment 3 again tests generalization of learning across reduction types and aims to replicate and extend the results from Experiment 1 on this issue.

Method

Participants. Sixty Dutch participants of the Max Planck Institute's subject pool, none of whom had participated in the previous experiments, took part for a small remuneration. All reported normal hearing and normal or corrected-to-normal vision.

Design. The design was very similar to the one of Experiment 1, except for changes in the exposure phase. Predictable exposure sentences were created for new potentially reduced /b/-words and vowel-deleted words which served as target words in an eye-tracking paradigm. That is, participants had to click on the orthographic form of these words while their eye movements were recorded. The test phase was the same as in Experiment 1. Due to the changes in the exposure phase, the targets in the test phase were new to participants and not repeated as in Experiments 1 and 2 (see Table 4).

Table 4. Experimental design and types of stimuli in Experiment 3. The potentially reduced /b/-initial words and the vowel-deleted words of the exposure phase were not repeated in the test phase.

	Condition	Canonical word-form	Segmental reduction group	Syllabic reduction group	Control group
			/b/ → [m]	full V deletion	no reduction
Exposure phase	Experimental	/bɑndit/	[mɑndit]	[bɑndit]	[bɑndit]
	Unreduced	/mɑtros/	[mɑtros]	[mɑtros]	[mɑtros]
	Experimental	/kanal/	[kanal]	[knal]	[kanal]
	Unreduced	/knɔflok/	[knɔflok]	[knɔflok]	[knɔflok]
Test phase	Experimental	/bɪndərɛɪ/		[mɪndərɛɪ]	
	Unreduced	/murɑs/		[murɑs]	
	Experimental	/parat/		[prat]	
	Unreduced	/xlɑns/		[xlɑns]	

Materials. The exposure sentences for the /m/-words and the CC-words were the same as in Experiment 1. The exposure sentences for the potentially reduced /b/-words and vowel-deleted words were constructed anew. These critical words appeared again towards the end of the sentences and were predicted by the semantic context.

For the selection of the 24 exposure /b/-targets and the 24 exposure vowel-targets, the same constraints applied as for the respective targets of the test phase. The criteria for the selection of their "competitors" were less strict. These only overlapped in the initial consonantal part for reduced forms, but were additionally matched on word class (e.g., *bandiet* 'bandit' would be reduced to [mɑndit] and would compete for recognition with [mirakəl] *mirakel* 'miracle'; *kanaal* 'canal' would be reduced to [knal] and would compete with [knɛxt] *knecht* 'servant'). The materials for the test phase were taken from Experiment 1.

Stimulus construction. The new exposure sentences were recorded by the same female Dutch speaker as in Experiments 1 and 2. The reduced stimuli were created as described in Experiment 1.

Procedure. The procedure was similar to Experiment 1 except for changes in the exposure phase, in which participants had to click on the potentially reduced word that was predictable from the sentence context. The computer display always showed a /b/-word, a /m/-word, a vowel-word and a consonant-cluster word on the screen. Exposure and test displays differed only in the phonological similarity of target and competitor words which were more similar in the test phase (e.g., exposure trial: *bandiet* vs. *mirakel*; test trial: *binderij* vs. *minderjarig*). An experimental session took approximately 25 minutes.

Results

Exclusion criteria. Trials were excluded based on the same criteria as used in Experiment 1 and 2. Due to fixations outside of the screen, 2.6% of the trials were removed. Another 0.6% were discarded due to failure to click on the target or the potentially confusable competitor. Fifteen trials (0.3%) in the exposure phase (with residuals either below -1100 ms or above 2500 ms) and 29 trials (0.5%) in the test phase (with residuals either below -1700 ms or above 2800 ms) were considered to be RT outliers, and hence excluded. For the eye-tracking results, the data of the participants' right eye were analyzed.

An overview of the accuracy data in the exposure and test phase can be found in Table 5. In the exposure phase, practically no errors were made. In the test phase, we observe again a high percentage of errors for the vowel-deleted words made by all three groups. Table 6 displays the descriptive statistics for the RT data in the exposure and test phases. All three groups took longer to give a click response in the test phase (where the sentence context was neutral and the words on the screen were quite similar to each other) than in the exposure phase (where they could use the sentence context to predict the target word). The negative minima for RTs in the exposure phase confirm that the target words in this phase were indeed predictable, as some participants responded even before target onset.

Table 5. Accuracy data of the exposure and test phases of Experiment 3.

% click responses		Segmental reduction group			Syllabic reduction group			No reduction group		
		Target	Comp.	Distr.	Target	Comp.	Distr.	Target	Comp.	Distr.
Exposure	/b/-word	99.4	0.0	0.0	99.4	0.0	0.0	99.6	0.0	0.0
	CVC-word	99.6	0.0	0.0	99.4	0.0	0.4	99.8	0.0	0.0
Test	/b/-word	95.8	3.1	0.0	95.2	4.2	0.0	92.3	6.7	0.2
	CVC-word	75.8	22.7	0.0	83.3	16.5	0.0	79.0	20.8	0.0

Table 6. RT in ms in the exposure and test phases of Experiment 3 for clicks on targets and competitors.

	Segmental reduction group		Syllabic reduction group		No reduction group	
RT in ms	Exposure	Test	Exposure	Test	Exposure	Test
mean	1280	1754	1429	1876	1418	1889
sd	610	736	761	947	644	955
min	-40	554	-40	430	-39	677
max	7374	8015	7060	10854	6419	9599

Exposure Phase. The /b/-words were reduced only for the segmental reduction group and the vowel-words were reduced only for the syllabic reduction group. There were virtually no errors (see Table 5). Moreover, in contrast to previous results (Chapter 3), there was no consistent effect of reduction, neither in reaction times nor in the eye-tracking data (see Figures 10, 11, and 12; the main effect of Group was not significant for the groups who heard reduced forms indicating that they did not have more difficulties in recognizing the targets than the control group). The data from the exposure phase reflect that the target words were predictable, as participants in all groups already showed a preference for the target before it was mentioned (see the time window from -200 ms to 0 ms in Figures 11 and 12). Apparently, the words in the exposure phase were recognized efficiently whether they were reduced or not.

Figure 10. Mean RTs and SEs in the exposure phase of Experiment 3.

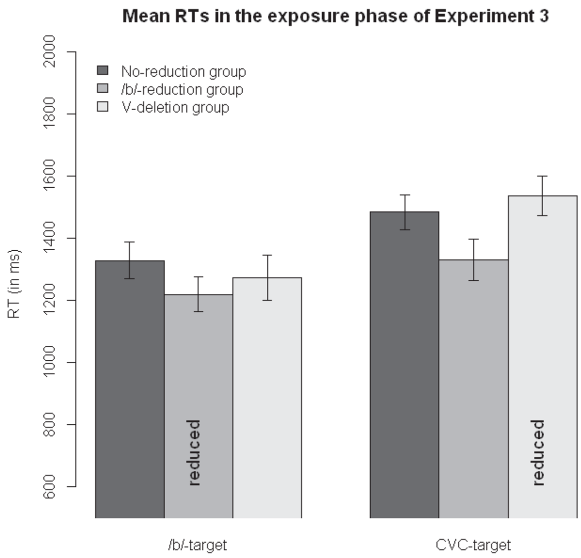
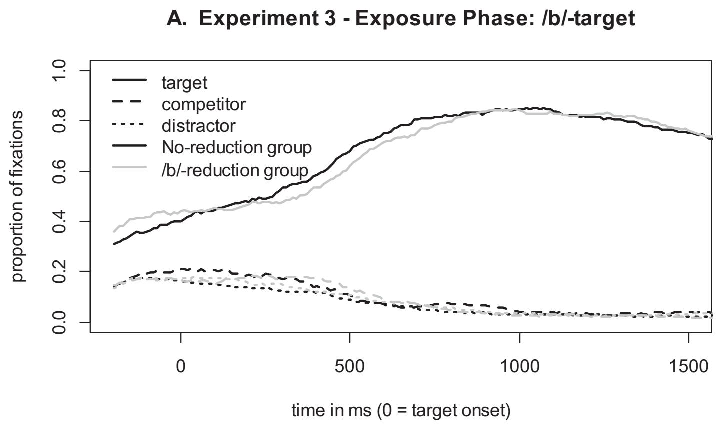


Figure 11. Proportion of fixations in the segmental reduction condition for the exposure phase of Experiment 3. The /b/-words were reduced for the segmental reduction group (in grey, Panel A), but unreduced for the control group (in black) and for the syllabic reduction group (in grey, Panel B).



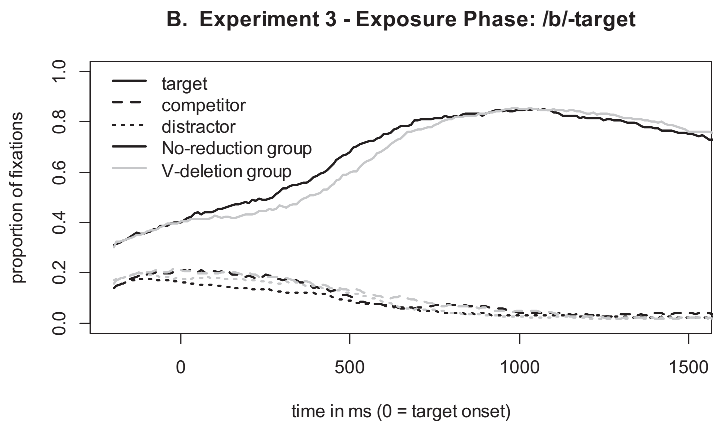
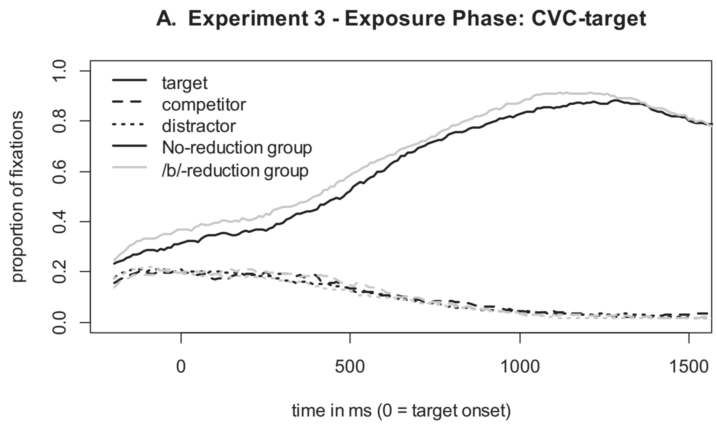
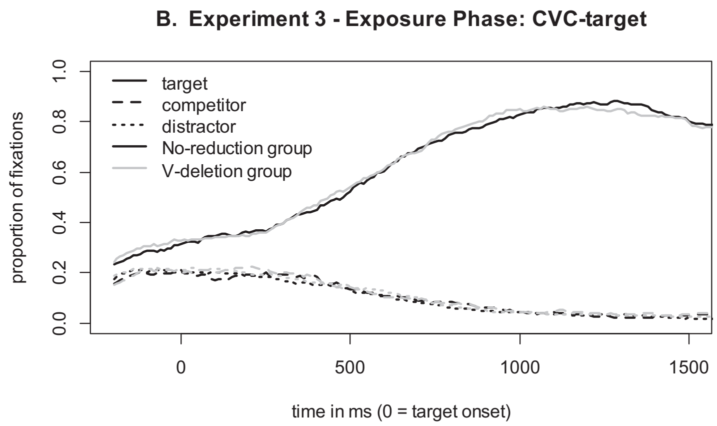


Figure 12. Proportion of fixations in the syllabic reduction condition for the exposure phase of Experiment 3. The vowel in the CVC-words was deleted for the syllabic reduction group (in grey, Panel B), but was present for the segmental reduction group (in grey, Panel A) and the control group (in black).

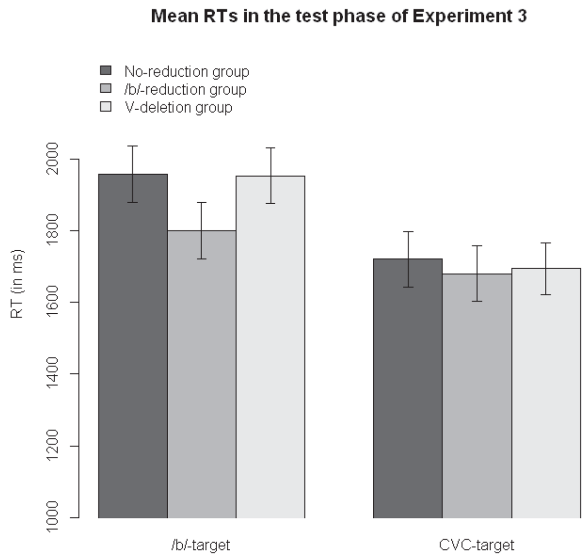




Test phase. ***Reaction time data.*** Figure 13 displays the mean RTs and SEs of all three groups for the reduced /b/-words and the vowel-deleted words in the test phase of Experiment 3. The segmental reduction groups seems to respond somewhat faster to the reduced /b/-words, while all three groups seem to respond about equally fast to the vowel-deleted words.

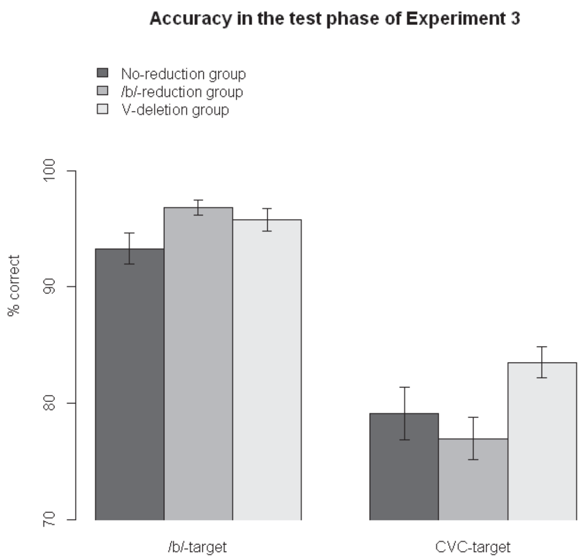
Statistical analyses did not show a main effect of Group—neither in the segmental reduction condition nor in the syllabic reduction condition—indicating that all groups responded equally fast to both types of target words. That is, the groups experienced with reduced forms did not respond faster than the less experienced control group. We thus did not observe any adaptation in the RT data.

Figure 13. Mean RTs and SEs in the test phase of Experiment 3.



Accuracy data. The accuracy data in terms of percentage correct responses and their SEs of all groups can be found in Figure 14. Both experimental groups seem to give more accurate responses to reduced /b/-words than the control group. For the vowel-deleted words, only the syllabic reduction group seems to respond more accurately than the control group. This, however, was not confirmed by statistical analyses. The main effect of Group was not significant either in the segmental reduction condition or in the syllabic reduction condition. That is, neither of the experimental groups gave more correct answers to the reduced targets than the control group. We thus did not observe any adaptation effects in the accuracy data.

Figure 14. Accuracy in % correct click responses and SEs in the test phase of Experiment 3.

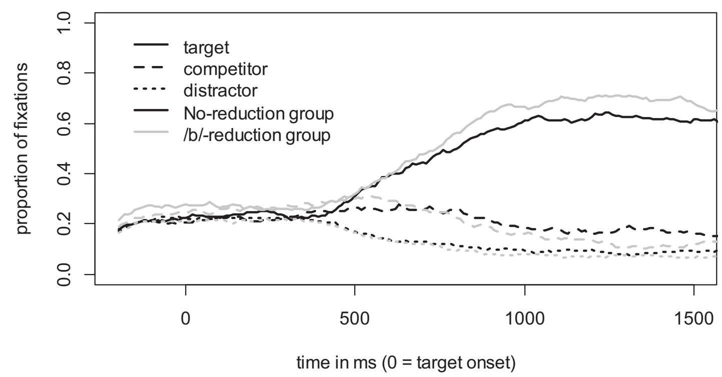


Eye movement data. Figure 15 displays the eye movement pattern of the two experimental groups plotted against the patterns of the control group (in black) for the segmental reduction condition. Both experimental groups show a greater preference for the target over the competitor for the reduced /b/-words than the control group, descriptively from around 700ms onwards (when the grey lines diverge from the black lines). This difference is bigger for the segmental reduction group (Panel A).

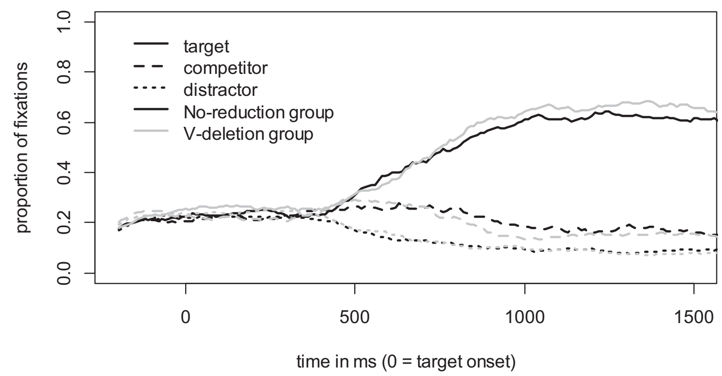
The main effect of Group reached significance only for the segmental reduction group ($b_{\text{Segmental reduction Group}} = 0.99$, $t = 2.4$, $p < .05$) from 1100 ms onwards. That is, the segmental reduction group, but not the syllabic reduction group, outperformed the control group on the reduced /b/-words. We thus observed a within-reduction-type generalization effect in the segmental reduction condition.

Figure 15. Proportion of fixations in the segmental reduction condition for the test phase of Experiment 3.

A. Experiment 3 - Test Phase: /b/-target



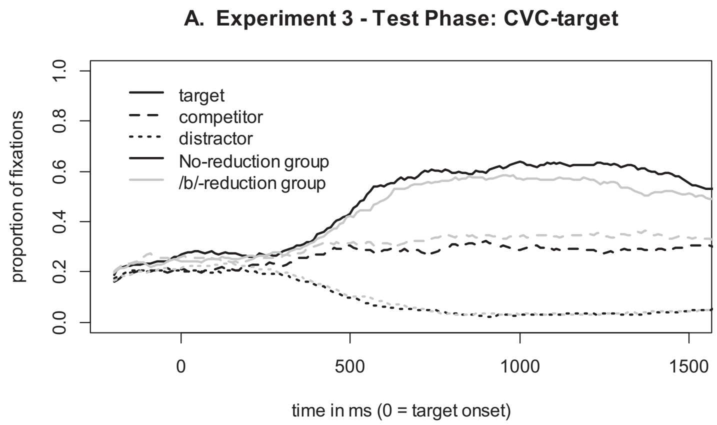
B. Experiment 3 - Test Phase: /b/-target

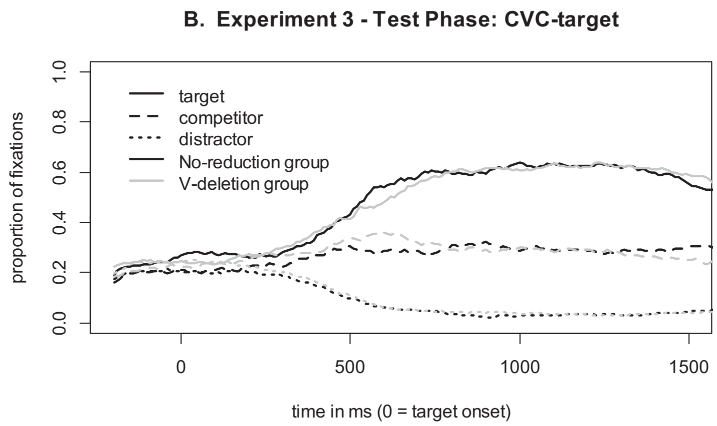


The corresponding eye-movement data for the syllabic reduction condition are displayed in Figure 16. The segmental reduction group (in grey, Panel A) shows a smaller target-competitor preference for the CVC-targets than the control group, descriptively from 500 ms onwards. The syllabic reduction group (in grey, Panel B) shows a similar pattern from 500 ms - 700 ms after target onset.

Statistical analyses showed a marginally significant main effect of Group only for the segmental reduction group ($b_{\text{Segmental reduction group}}$ = -0.93, t = -2.0, p = .06) in the time window from 1200 ms - 1500 ms. That is, the segmental reduction group but not the syllabic reduction group had a significantly smaller target preference for the vowel-deleted words than the control group. We thus did not observe a learning effect for the syllabic reduction group and found a marginal inhibitory effect for the segmental reduction group. Participants in the latter group seem to be hindered by their prior exposure to another reduction type.

Figure 16. Proportion of fixations in the syllabic reduction condition for the test phase of Experiment 3.





Discussion

The aim of Experiment 3 was to test whether learning about reductions can generalize within and across reduction types. In the exposure phase, listeners were provided with predictive sentence contexts and with orthographic information about the critical words, as they saw the orthographic form of the potentially reduced word on the computer screen. The results from the exposure phase did not show any effects of reduction. That is, neither the segmental reduction group nor the syllabic reduction group were slowed down or had a smaller target preference when hearing reduced forms. This is very likely due to the predictive sentence context. Participants were already expecting the target and looking at it before it was actually mentioned. Hearing it then in reduced form did not disturb the recognition process any more. Note that these data apparently are in contrast with the data of Brouwer et al. (2013), who found that even predictable words suffer from reduction costs. The difference, however, might be due to the stimulus material, with our material being constructed to allow prediction of the target word, while Brouwer et al. used materials from a speech corpus. For reduced words which were particularly predictable, they also observed less reduction costs.

In the test phase, we found clear evidence for generalization of learning within reduction type for the segmental reduction group in the eye-tracking data. No such generalization effect was found for the syllabic reduction group. Contrary to the word-specific learning effects found in Experiment 1 and 2, the within-reduction-type generalizations were stronger for /b/-reductions than for vowel-deletions.

As for generalization of learning across reduction types, we did not replicate the transfer of learning from vowel-deletions to /b/-reductions for the syllabic reduction group found in Experiment 1. There was a trend going in this direction though (see Figure 15B). However, we replicated the marginal inhibitory effect of the segmental reduction group found in Experiment 2. That is, the segmental reduction group did not benefit from its exposure to /b/-reductions and had instead slightly greater problems in recognizing vowel-deletions than the no-reduction control group.

General Discussion

The present study investigated whether and how listeners can adapt when they encounter reduced word forms. In the introduction, we argued for a continuum of possible adaptation mechanisms that are more or less general. At the specific end, listeners may only adapt to exactly the same words. A more general adaptation would allow generalization to other words of the same or a similar reduction type. Experiment 1 and 2 tested learning effects for repeated segmental and non-morphemic syllabic reductions. Experiment 3 determined whether these learning effects were word-specific by testing whether learning about these reductions generalizes to new words of the same reduction type (within-reduction-type generalization). All three experiments investigated whether experience with one reduction type helps the listener in dealing with another reduction type (across-reduction-type generalization).

Experiments 1 and 2 showed clear evidence of learning for repeated vowel-deletions but, surprisingly, far less so for repeated /b/-reductions. In contrast, Experiment 3 revealed a strong within-reduction-type generalization effect for the /b/-reductions that was not found for the vowel-deletions. In Experiments 2 and 3, the segmental reduction group further showed a marginal inhibitory effect; they had greater difficulties than the control group dealing with unfamiliar vowel-deletions. Another pattern that was consistently observed (even though not always significant) was that the syllabic reduction group made less errors for both the same and other vowel-deleted words (see Figures 3, 7, and 14, focusing on the CVC-targets, e.g., *paraat* produced as *p'raat*). Next to this reduction-specific adaptation, this group also showed generalization of learning across reduction types (from vowel-deletions to /b/-reductions). This generalization effect, however, could not always be found: It was absent in Experiment 2 where task demands in the exposure phase were low and the predictability of the reduced word was high. It was present in Experiment 1, where task demands in the exposure phase were high, but the predictability of the reduced word was low. Finally, a trend was observed again in Experiment 3, where both task demands and the predictability of the reduced word in the exposure phase were high.

The results of Experiment 3 shed further light on the learning effects found in Experiment 1 and 2. For the segmental reductions, strong generalization of learning to new reduced /b/-words was observed. This suggests that, for the /b/-reductions investigated here, recognition predominantly occurs via abstraction rules. It is therefore likely that abstraction processes also play a role in the recognition of repeated reduced /b/-words. The learning effect found for repeated reduced /b/-words in Experiment 1 thus is very likely not a word-specific adaptation. For the vowel-deletions, no generalization of learning to other vowel-deleted words was observed in Experiment 3. The adaptation effects for repeated vowel-deleted words found in Experiment 1 and 2 are therefore very likely due to storage of these reduced forms and hence are word-specific. Similarly, Hanique, Ernestus, and Schuppler (2013) claim that, if the absence of schwa in the prefix of Dutch past participles is due to categorical processes, these schwa-deleted forms are stored in the mental lexicon.

Episodic storage is not only useful if a listener encounters a reduced word for the first time, but may also help to build up abstraction rules for later generalization of learning to other words that show the same reduction pattern. It is therefore surprising that we did not find any benefit for repeated reduced /b/-words in Experiment 2, while we did find a benefit for repeated vowel-deleted words under the same circumstances. Furthermore, although small, such a benefit was found for repeated reduced /b/-words in Experiment 1, where participants were involved in a more active task, but where the reduced /b/-words were hardly predictable. One possible explanation for these findings is based on the difference in saliency between the two reduction types. In the vowel-deletions, an entire segment is completely deleted, whereas in the /b/-reductions the segment is only weakened. The vowel-deletions are thus more striking than the /b/-reductions and potentially are therefore less susceptible to experimental manipulations. Apparently, manipulating the preceding context to make the reduced /b/-words more predictable was not enough to draw participants' attention to that reduction type, while giving listeners a more active task might have achieved this. Learning about reductions might thus only occur if the reduction type is (made) salient enough. Note that in Experiment 3, where learning for /b/-reductions was found, listeners saw the orthographic form of the reduced /b/-words on the screen already in the exposure phase. This may have boosted the learning effect.

The within-reduction-type generalization effect found for new reduced /b/-words in Experiment 3 supports the assumption of an abstractionist mode of lexical access. For the vowel-deletions, only a hint of this generalization effect was observed (in the accuracy data). An important difference between /b/-reductions and vowel-deletions that could explain this discrepancy is input consistency. In the /b/-initial words that were to be reduced, the /b/ was always followed by a vowel and a nasal. The structure of the CVC-words was less consistent: The first consonant could be /k, x, p/, the vowel to be deleted was variable and the second consonant was either a liquid or /n/. The phonological context surrounding the reduced segment and the reduced segment itself varied thus more in the vowel-deletions than in the /b/-reductions. This input variability for vowel-deletions may have been too high for the successful generation of an abstract mapping rule. This very likely restricts generalized learning about syllabic reductions to morphemes that show a high frequency of occurrence across words.

There hence seems to be evidence for two types of adaptation: word-specific adaptation to inconsistent phonological patterns and word non-specific adaptation to consistent patterns. More general learning effects, if observed at all, were marginal. This already suggests that it is hard to apply the knowledge of one reduction type to another in case the two reduction types differ substantially. Nevertheless, we observed such a non-specific adjustment to reductions for the syllabic reduction group. Listeners in this group showed a greater tolerance to /b/-reductions than the control group. Possible factors that likely play a role in this uni-directional facilitative effect are input variability and degree of reduction. These two factors, however, are (necessarily) confounded in the present study. The vowel-deletions are both more variable in their segmental structure and more severely reduced than the /b/-reductions. Similar conditions were present in the study by Brouwer, Mitterer and Huettig (2012) who found similar facilitative effects. They reported that listeners penalized acoustic mismatches between input and canonical form less strongly when listening to (strongly and therefore not regularly) reduced speech.

Instead of also observing facilitation for the segmental reduction group in dealing with vowel-deletions, we found marginal inhibitory effects. After having been exposed to consistently reduced /b/-words, the segmental reduction group did worse on the more strongly reduced vowel-words than the control group. It might thus be that learning about reduction can only generalize to other reduction types that are of the same or a lesser degree of reduction, but not to reduction types that show a higher degree of reduction. Another possibility is that the vowel-deletions differed in too many ways from the /b/-reductions so that it was not possible to adjust the abstract mapping rule for /b/-reductions to accommodate the variable vowel-deletions.

But why did the segmental reduction group actually differ from the control group in dealing with vowel-deletions? It might be the case that participants in the segmental reduction group expected the speaker to only produce reductions in a consistent way and of a specific degree (weakening of a segment). This might have biased them against other types of variability and the greater deviation from the canonical form that they encountered in the test phase. The control group, in contrast, had not heard any reductions in the exposure phase. In the subsequent test phase, participants in that group suddenly had to deal with many and various reduced forms. As they could not have built up abstract mapping rules, they probably resorted to flexible, non-specific adjustments, as those observed by Brouwer et al. (2012).

Finally, the syllabic reduction group was already used to dealing with variable reduced forms. Participants in this group could therefore handle a consistent and less severe reduction type. How well listeners can handle new reduced forms of a different reduction type might thus also depend on listeners' expectations about a speaker's reduction style and, based on that, on the adaptation mechanisms already in use (specific abstraction rules vs. fast perceptual but non-specific adjustments).

What does this series of eye-tracking experiments tell us about possible constraints and the time course of learning about reductions? Apparently, the reduced forms have to be noticeable, as learning effects were found for less salient reduction types only if the reduced words appeared in orthographic form on the screen (Experiment 3) or if the listener was actively involved in the task (Experiment 1), whereas this was not necessary for salient reduction types. Interestingly, the generalization effects across reduction types varied in strength across experiments, which suggests that at least some part of learning is susceptible to our experimental manipulations. Attention as measured by task involvement (Experiment 1) seems to be of greater importance than predictability (Experiment 2) in dealing with reductions. However, the combination of these two factors (Experiment 3) yielded only a trend in the expected direction.

Moreover, the time course suggests that the point in time when learning about reductions takes effect may depend on the specificity of the learning process. Facilitative and inhibitory generalization effects across reduction types, which are likely not specific to any segments or words in our study, were observed early throughout the study (from 200ms and 300ms after target onset respectively). Also the inhibitory effect in Experiment 3 emerged early (around 500 ms after target onset), but reached marginal significance only late (at 1200 ms). In contrast, the effect for generalization within reduction type in Experiment 3 was quite late (starting at 1100 ms after target onset). The word-specific effect found in Experiment 1 was equally late. The former may be explained with the kind of mapping procedure participants have to apply. Listeners learned that this particular speaker was likely to pronounce a /b/ as an [m] and hence that an existing sound ([m]) mapped onto two categories for that speaker (/m/ and /b/). Their perception of an [m] might therefore have shifted from judging it as /m/ in most cases to judging it as /m/ in 80% and as /b/ in 20% of the cases. With this kind of learning, an initial signal-driven hypothesis strongly favors the canonical form, and only when later-arriving segments rule that form out can the learning take effect. Therefore, as soon as listeners receive evidence that a particular sound can map onto more than one category, the rule-based learning process likely needs more time to take effect. A similar reasoning can be applied to word-specific learning. At some point in time, the activation of *Parijs* 'Paris' has to win over the activation of *prijs* 'price' when hearing the reduced speech input *P'rijs*. Initially, the activation of *prijs* is likely to be stronger as this meaning is encountered much more frequently. Speaker-specific information (e.g., on the tendency of this speaker to reduced words like *Parijs*) has then to kick in and shift the weights in favor of the candidate *Parijs*. This may take some time and not happen immediately.

Conclusions

The present study provided evidence that listeners use a wide variety of adaptation mechanisms when dealing with reduced forms. Word-specific learning effects showed that reduced forms were stored as such in the mental lexicon. If possible, that is, if the input provided sufficient consistency, abstraction rules were generated based on the reduced speech input and applied to new reduced words. In the setting of the present study, this was only successful for new words of the same reduction type. If the input was too variable, listeners also showed perceptual flexibility and were able to deal with various reduction types. The interplay of abstraction processes and perceptual adjustments may come at a cost if abstract mapping rules are already in place. The perceptual system might then not be flexible enough to allow rapid accommodation of inconsistent reductions. To conclude, both episodic and abstractionist modes of lexical access, as well as perceptual flexibility, play a role in recognizing reduced word forms.

Appendix

Table A1. Target and competitor words in the segmental reduction condition of the test phase in Experiments 1, 2, and 3 with their word frequency per million according to SUBTLEX-NL (Keuleers, Brysbaert, & New, 2010).

Test phase: segmental reduction condition					
Target	English Translation	Word Frequency	Competitor	English Translation	Word Frequency
banaal	banal	1.2	manuaal	manual	0.0
banaan	banana	5.3	mangaan	manganese	0.1
banier	banner	0.5	manie	mania	0.5
banket	banquet	3.0	mangrove	mangrove	0.1
bankier	banker	3.7	mankeren	to be wrong	1.0
bemachtigen	to get hold of	3.3	metallic	metallic	0.2
bemesten	to manure	0.1	memento	memento	0.1
bemoeien	to meddle	12.1	miljoenen	millions	30.4
benaming	name	0.6	mekaar	each other	16.2
benard	awkward	0.0	miljard	billion	12.2
benauwen	to oppress	0.3	mevrouwen	women	0.1
beneden	below	188.0	meneer	gentleman	518.4
benedictijn	Benedictine	0.0	menigvuldig	manifold	0.0
benedijen	to bless	0.0	menageren	to moderate	0.0
benoemen	to appoint	2.5	meloenen	melons	1.6
benoorden	north of	0.0	mesjokke	crazy	0.4
benutten	to utilize	1.6	mejuffrouw	Miss	2.2
benzine	gas	24.7	menslievend	charitable	0.2
binair	binary	0.3	mineur	minor	0.4
binderij	bookbinder	0.0	minderjarig	underage	2.9
binnenkort	soon	45.5	minnenswaard	lovable	0.0
binomisch	binomial	0.0	minoriteit	minority	0.0
bonbon	chocolate	0.3	mondain	faishonable	0.2
bonjour	good day	2.6	montuur	frame	0.4
Average		12.3	Average		24.5

Table A2. Target and competitor words in the syllabic reduction condition of the test phase in Experiments 1, 2, and 3 with their word frequency per million according to SUBTLEX-NL (Keuleers, et al., 2010).

Test phase: syllabic reduction condition					
Target	English Translation	Word Frequency	Competitor	English Translation	Word Frequency
charisma	charisma	0.8	chrisma	chrism	0.0
correct	correct	15.3	krek	just	0.1
coulisse	wing	0.0	klissen	to get entangled	0.1
curator	guardian	1.1	krater	crater	2.1
galei	galley	0.5	glei	thatch	0.0
galop	gallop	0.5	glos	gloss	0.0
genoom	genome	0.3	gnoom	gnome	0.3
gering	small	1.0	grind	gravel	0.7
kaneel	cinnamon	1.5	kneep	pinch	2.3
kanon	gun	6.2	knol	tuber	0.9
karaat	carat	1.1	kraak	crack	2.0
karaf	carafe	0.2	kras	scratch	2.1
karos	coach	0.1	krols	on heat	0.0
koliek	colic	0.2	kliek	clique	1.0
kolom	column	0.9	klomp	clog	0.9
kolonie	colony	6.4	clonen	to clone	0.1
kolos	colossus	0.4	klos	bobbin	3.4
konijn	rabbit	18.9	knijp	stuck	4.7
koran	Koran	2.0	kraan	tap	6.4
paraat	ready	6.2	praat	talk	241.8
Parijs	Paris	55.4	prijs	price	86.6
piloot	pilot	30.1	plooi	fold	0.6
polijsten	to polish	0.4	pleister	plaster	2.4
puree	puree	4.5	pree	pocket money	0.0
Average		6.4	Average		14.9

Table A3. Target words in the segmental and syllabic reduction condition of the exposure phase in Experiment 3 with their word frequency per million according to SUBTLEX-NL (Keuleers, et al., 2010).

Exposure phase: potentially reduced target words					
/b/-target	English Translation	Word Frequency	CVC-target	English Translation	Word Frequency
bandiet	bandit	4.7	caleidoscoop	kaleidoscope	0.2
bemanning	crew	20.4	canard	newspaper hoax	0.1
bemantelen	to cloak	0.0	Canossa	Canossa	0.0
bemerken	to notice	0.2	cholerisch	choleric	0.0
bemeten	sized	0.0	collaps	collapse	0.0
bemiddeling	mediation	0.2	college	lecture	13.3
beminnen	to love	1.8	collisie	collision	0.0
bemoederen	to mother	0.2	courage	courage	0.3
bemoedigen	to encourage	0.2	courant	newspaper	0.2
bemorsen	to soil	0.0	coureur	race-driver	2.0
bemost	bemoost	0.0	galactisch	galactic	0.2
benadelen	to harm	0.2	garantie	guarantee	6.1
benadering	approach	2.8	gareel	harness	1.2
benadrukken	to emphasize	2.2	generisch	generic	0.0
Benelux	Benelux	0.0	kanaal	channel	13.9
benemen	to take away	0.2	kinine	quinine	0.5
benepen	small-minded	0.2	kolere	cholera	6.6
benevens	besides	0.0	koraalrif	coral reef	0.2
benieuwen	to arouse curiosity	0.0	koraalrood	coral red	0.0
benijden	to envy	1.1	paleis	palace	15.5
benodigd	necessary	0.3	parochie	parish	2.5
benul	notion	2.9	piraat	pirate	6.4
bombardement	bombardment	1.8	polemisch	polemic	0.0
bombastisch	bombastic	0.0	politie	police	346.2
Average		1.6			17.3

Familiarity with regional variation in speaking style is necessary for effortless recognition of morphologically complex words in casual speech

Chapter 5

Poellmann, K., McQueen, J. M., & Mitterer, H. (submitted). Familiarity with regional variation in speaking style is necessary for effortless recognition of morphologically complex words in casual speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Abstract

An ERP experiment tested whether familiarity with regional variation in speaking style influences how listeners recognize familiar and unfamiliar pronunciations of words in casual speech. N200 and N400 ERP components measured processing difficulties when listeners were confronted with reduced pronunciations of morphologically complex words. These included one regionally constrained reduction type (the prefix *ge*- reduced to [g]/[k], as occurs in Southern German and Austrian accents) and one that is relatively unfamiliar throughout Germany and Austria (the prefix *ver*- reduced to [f]). Three experimental groups (Northern Germans, migrant Northern Germans living in the South and Southern Germans) heard sentences containing these reduced forms. A control group heard sentences with only unreduced forms. For *ge*-reductions, the results showed a benefit of long-term exposure for the Southern listeners (neither an N200 nor an N400 effect) and fast perceptual adjustments for the Northern and migrant listeners (an N200, but no N400 effect). For *ver*-reductions, the Southerners showed perceptual flexibility (an N200, but no N400 effect), whereas the

Northerners and especially the migrants had persistent processing difficulties (N200 and N400 effects). Reductions resulting in illegal phonotactic sequences caused early processing difficulties, but these were overcome fast unlike the continuing processing costs caused by reductions leading to legal sequences. The results suggest that familiarity with specific reduction styles influences processing. For accent-standard pronunciations, accent speakers develop representations of those variant words, while speakers from elsewhere recognize such variants through rapid perceptual adjustments. Furthermore, experience with accent-standard pronunciations gives speakers enhanced flexibility in recognizing non-standard pronunciations.

Introduction

In casual conversations, speakers do not necessarily pronounce every word carefully. Instead, segments and even syllables frequently get reduced (Ernestus, 2000; Johnson, 2004; Patterson, LoCasto, & Connine, 2003). The tendency of a speaker to reduce words may depend on the region s/he grew up in (Keune, Ernestus, van Hout, & Baayen, 2005). It is, for example, very likely that a Southern German or Austrian speaker would leave out the schwa in the prefix *ge-* /gə/, so that the past participle *geliebt* /gəliːpt/ 'loved' would be pronounced as [gliːpt]. A Northern German speaker is less likely to produce such reductions. Both speakers, however, may shorten the prefix *ver-* /fɛɐ/ to [fɐ], but probably not frequently to [f]. That is, they may pronounce the past participle *verliebt* /fɛɐliːpt/ '(fallen) in love' as [fɐliːpt], but rarely as [fliːpt]. The present study investigates how familiarity with certain reduction types influences processing of reduced words. German listeners who are familiar with a regionally constrained reduction type (reduction of the prefix *ge-* to [g]/[k]) may process these reductions differently from German listeners who are unfamiliar with that kind of reduction. Moreover, the present study tests whether experience with certain reduction types is necessary for adaptation to take place. That is, can listeners adapt to reductions with which they are relatively unfamiliar (reduction of the prefix *ver-* to [f])?

Languages differ in the frequency and type of reduction phenomena that they show. In Dutch, for example, affix reductions, such as the reduction of the inflectional prefix *ge-* /xə/, occur frequently (see the Corpus of Spoken Dutch, Oostdijk, 2000; Pluymaekers, Ernestus, & Baayen, 2005), whereas reductions of the inflectional prefix *ge-* /gə/ can hardly be found in Standard German (see, for example, the Kiel Corpus of Spontaneous Speech, IPdS, 1994). But other reductions, some quite extreme, are nevertheless documented in the Kiel Corpus of Spontaneous Speech. One can thus listen to *Stunden* /ʃtʊndən/ 'hours' pronounced as [ʃtʊn], or *Montag* /moːntaːk/ 'Monday' realized as [mõtaː].

This difference in the frequency and type of reduction phenomena is not limited to comparisons across languages, but may also be found in dialects or regional variants of a language. In Southern Germany and Austria, it is very common to reduce the prefix *ge-* /gə/. It is indeed so common that it can even be found it in written form. The *Süddeutsche Zeitung,* for example, headlined on February 14th 2013 (Crone & Fuchs, 2013): "Scherzerl gmacht, keiner glacht" (which roughly translates as 'joke been made, no one laughed'), where the vocalic part of the prefix has been omitted in the past participles *gemacht* 'made' and *gelacht* 'laughed'. Such reductions are not documented for the Northern German variety.

Usually, neither speakers nor listeners are aware that they produce and encounter reductions frequently (Ernestus & Warner, 2011; Kemps, Ernestus, Schreuder, & Baayen, 2004). However, as soon as reductions occur as characteristics of regional variation, they are much more prominent, as demonstrated in the newspaper example.

As the frequency and type of reduction that a listener encounters may depend on the origin of the speaker (Keune, et al., 2005), it is important to know how listeners deal with regional variation. Sumner and Samuel (2009) investigated the effect of experience on the comprehension of dialect variants. They tested three groups varying in their experience with r-dropping in *–er* final words that is common in the New York City area. Participants in the General American group were born and raised outside the New York City region and unfamiliar with r-less dialects. The two other groups were born and raised in the New York City area, but differed in whether they actively produced r-less variants (Overt-NYC group) or not (Covert-NYC group). Participants in the Overt-NYC group were third-generation New York City area residents and heard r-less variants at home. Participants in the Covert-NYC group were first- or second-generation NYC residents whose parents originally came from other states or other countries and who heard r-less variants mostly from peers. The r-less variants posed problems for the General American group. The two NYC groups had no difficulties in understanding r-less variants and showed short-term priming effects for these words. Participants in the Covert-NYC group, however, differed from participants in the Overt-NYC group in that the former did not show long-term priming effects for r-less variants. Therefore, Sumner and Samuel described the Covert-NYC participants as “fluent listeners” who exhibit immediate perceptual abilities to recognize r-less word forms that they however do not produce themselves and that they do not store as such in their mental lexicon.

Floccia, Goslin, Girard, and Konopczynski (2006) examined the time course of comprehension disruption caused by regional and foreign accents in a series of lexical decision experiments. They found that after initial word identification delays, which were smaller for regional accents than for foreign accents, participants achieved full adaptation for regional accented speech whereas their processing deficit did not reduce over time for foreign accented speech.

The present study extends those of Sumner and Samuel (2009) and Floccia, et al. (2006) insofar as it examines the time course for processing regional variants in more detail by measuring event-related brain potentials. Moreover, we also advance research on adaptation processes by investigating whether listeners can adapt to an uncommon reduction type that is related to more common reductions (i.e., the common reduction of the prefix *ver-* to [fɐ] is taken one step further and reduced to [f]).

We investigated the means by which listeners adapt to reduced speech. They use a variety of sources of information. To recover from reductions such as schwa-deletions, for instance, listeners can exploit fine phonetic detail. Manuel (1992) reported that listeners are able to distinguish a reduced form [spɔːt] of 'support' from the unreduced form [spɔːt] 'sport'. Phonotactic constraints also help listeners to overcome schwa-deletions. In a cross-modal priming experiment, Spinelli and Gros-Balthazard (2007) did not observe processing costs for schwa-deleted word forms that resulted in illegal phonotactic sequences in French, whereas they did so for reduced word forms that resulted in legal phonotactic sequences. One way to explain this result is that the illegal sequence triggers repair processes that induce an epenthetic schwa in perception (e.g., Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999). Legal sequences, however, may activate all words that start with the legal sequence (e.g., a reduced form of *geliebt* [gliːpt] may activate words starting with /gl/, such as *Glück* 'luck'), which in turn hinders the activation of the intended word.

Listeners also make use of the sentence context to compensate for reduced words (Brouwer, Mitterer, & Huettig, 2013; Ernestus, Baayen, & Schreuder, 2002; Mitterer & McQueen, 2009). They use, for instance, probabilistic cues from the surrounding context to compensate for reductions (Mitterer & McQueen, 2009). In addition to the phonetic context, the syntactic/semantic context is beneficial for the listener, especially when dealing with highly reduced word forms (Ernestus, et al., 2002; Van de Ven, Tucker, & Ernestus, 2011). Prefix reductions were therefore implemented in the present study at the end of sentences with highly predictive semantic contexts.

As a phenomenon of stylistic variation, reductions are optional. Factors favoring the occurrence of reductions are for example fast speaking rate, high word predictability, high word frequency, and appropriate phonological context, such as unstressed position within a word (Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Warner & Tucker, 2011). Because they are unstressed, affixes are good candidates for undergoing reduction processes (Pluymaekers, et al., 2005).

Therefore, we investigated two types of prefix reduction (*ge*- and *ver*-), which were implemented in German past participles, that is, in the same word form of one word class. In German, the past participle is formed by adding the inflectional prefix *ge*- to the stem of the verb and attaching the suffix -*t* (if the verb is regular) or the suffix -*en* (if the verb is irregular). The past participle of the regular verb *lieben* 'to love' is therefore *geliebt* 'loved', while the past participle of the irregular verb *schwimmen* 'to swim' is *geschwommen* 'swum'. In irregular verbs, the vowel of the stem often changes when the past participle is formed. Importantly, the inflectional *ge*-prefix is not used for the past participle formation if the underlying verb already has a derivational prefix like *ver*-. Thus, the past participle of the regular verb *verlieben* 'to fall in love' is *verliebt* '(fallen) in love' without an additional *ge*-prefix. These past participles were presented to listeners in reduced form as highly predictable words at the end of sentences. The *ge*-prefixes were reduced to [g], if the first segment of the stem was voiced. That is, *geliebt* /gəliːpt/ 'loved' was reduced to [gliːpt]. If the first segment of the stem was unvoiced (as in *geschwommen* /gəʃvɔmən/ 'swum'), the prefix *ge*- was reduced to [k] (e.g., [kʃvɔmən]). The prefix *ver*- was always reduced to [f]. Note that, while *ver*-to-[f] reductions are rare in German, comparable reductions do occur, for instance, in Dutch. The infrequent occurrence of such forms in German is thus not due to some universal faithfulness constraints.

A major goal of the present study was to look at adaptation effects as listeners are or become more familiar with a certain type of reduction. This study was designed to investigate both short-term and long-term adaptation effects. Long-term adaptation effects were investigated by presenting regionally constrained *ge*-reductions to different listener groups: Northern Germans, Northern Germans living in the South (migrants), and Southern Germans. Short-term adaptation was investigated by examining whether the processing difficulties caused by reduced forms changed over the course of the experiment.

Within our experimental design, observing short-term adaptation, that is, adaptation to reductions within the course of the experiment, would be evidence for an abstract mode of lexical access, with abstraction taking place at a prelexical level. This is especially true if found for the *ver*-reductions. As they occur only rarely, participants are unlikely to have stored and accessed such pronunciation variants when recognizing these forms. As words were not repeated in this experiment, a facilitation in recognizing reduced words in the second half of the experiment (compared to the first half) would mean that listeners have built up a mapping procedure that restores the reduced vocalic parts of the two prefixes under investigation. This generalization of learning about sublexical units to new reduced words is something that cannot be achieved by a strictly episodic model (Cutler, Eisner, McQueen, & Norris, 2010; McQueen, Cutler, & Norris, 2006). In the episodic mode of lexical access, recognizing some reduced words should have no influence on recognizing other reduced words of the same reduction type, as the weights for specific entries cannot be adjusted for the realization of a prefix only and, importantly, they cannot be adjusted across words.

Observing long-term adaptation effects, that is, a group difference in processing the reduced variants with advantages for the group(s) familiar with a specific reduction type (*ge*-reductions), would indicate the existence of multiple (possibly abstract) representations. We expect that at least the Southern group will have representations for canonical and reduced variants of *ge*-words and therefore will show no difficulties in processing the *ge*-reductions. It seems unlikely that abstractionist mapping rules (e.g., [k] → *ge*-) could also account for an advantage of familiarity with a certain reduction type. To do so, the mapping rules, once acquired, would need to be susceptible to changes in experience with that reduction type and so would be strengthened, for instance, if listeners encounter certain reductions more often, and weakened if they encounter them only rarely.

Another point addressed in this study concerns the issue of phonotactic legality. The vocalic parts of the German prefixes *ge*- /ɡə/ and *ver*- /fɛɐ/ were deleted which led to illegal phonotactic sequences in half of the *ge*- and half of the *ver*-words. We asked whether reductions that result in phonotactically legal sequences (i.e., legal consonant clusters in onset position, such as *geliebt* > *gliebt, verliebt* > *fliebt*) elicit different ERPs from reductions that result in sequences which are phonotactically illegal in standard German (e.g., *geschwommen* > *kschwommen, vermisst* > *fmisst*). The literature provides the basis for opposing predictions. On the one hand, Spinelli and Gros-Balthazard (2007) found that reductions resulting in phonotactically illegal sequences cause less processing difficulties, which may be due to repair strategies invoked by the illegal sequences. On the other hand, it can be argued that legal sequences should be easier for the listener. Connine, Ranbom, and Patterson (2008) argued that variants are easier to process if they are frequent. Literature on the determinants of reduction (reviewed in Shockey, 2003) indicates that reductions occur more often if the resulting phoneme sequence is phonotactically legal. This would lead to the prediction that reductions resulting in legal sequences are more familiar and hence easier to process.

Note that some sequences that were classified as illegal for the purpose of this experiment (e.g., word-initial <gschw> as in *g(e)schwommen* 'swum' and <gm> as in *g(e)malt* 'painted') occur in place names in Southern Germany (e.g., *Gschwend, Schwäbisch Gmünd*). Moreover, all *ge*-reductions (including the ones resulting in what are characterized in standard German as illegal phonotactic sequences) actually do occur at least in Southern German speech, whereas the *ver*-reductions (legal and illegal) occur relatively rarely. It will thus be another point to examine whether this difference is reflected in participants' ERPs.

This is the first study to examine reductions with event-related brain potentials (ERPs). The aim is to compare the processing of the reduced (*ge*- and *ver*-) words to the processing of canonical (*be*-) words and nonwords. The use of ERPs has two advantages here. First of all, we can examine the processing of reductions without an additional meta-linguistic task. Moreover, we can see how these reduced forms are processed in time, as different ERP components arising at different latencies have been associated with difficulties in word processing.

An early component that might indicate processing difficulties for words that are reduced word-initially is the N200. This is a negative-going wave that peaks around 200 ms post-stimulus onset. It is said to reflect the operation of a mismatch detector, which is sensitive to perceptual novelty. It can be elicited when a word does not start in a way predicted by the listener on the basis of the preceding sentence context (Connolly & Phillips, 1994; Van den Brink, Brown, & Hagoort, 2001). The critical stimuli in the present study all occurred in positions where a past participle was expected. Past participles in German usually start with a CV syllable. If participants encounter a reduced past participle starting with a consonant cluster they might exhibit a more negative N200 than when hearing a canonical (unreduced) participle.

Another important component with respect to word processing difficulties is the N400 (for a review, see Kutas & Federmeier, 2011). This is a negative-going potential that peaks around 400 ms post-stimulus onset. Factors that influence the N400 amplitude are, for example, word frequency (with lower frequent words eliciting a larger, that is, more negative N400), neighborhood size (with larger neighborhood sizes eliciting larger N400s) and cloze probability (with less predictable words eliciting larger N400s). Importantly, N400 activity has also been elicited by pronounceable pseudowords. The processing of canonical words should show a moderate N400, while the nonwords were predicted to evoke a much more negative N400, as they usually do not occur, and hence are less expected and less predictable than canonical words. The question is then how reduced words are processed. If it is easy to access their meaning despite the reduction that they underwent, they should be processed like canonical forms. If the reduction process, however, was so severe that listeners do not recognize the originally intended word anymore, that reduced word will be as hard to integrate into the sentence context as a nonword and a large N400 should be observed. The amplitude of the N400 is thus a measure of processing ease. The more negative the amplitude is, the harder it is to process that word.

To summarize, the present study had five goals: We explored the time course of on-line recognition of prefix reductions, we examined the effect of familiarity on the processing of a reduction type that is standard in Southern German (*ge*-reduction), we investigated the effect of familiarity with one reduction type on processing another, unfamiliar reduction type, we tested whether listeners can adapt quickly (within the course of the experiment) to a relatively unfamiliar kind of reduction (*ver*-reduction), and we investigated a possible difference in processing of reduced word-forms that start with a phonotactically illegal sequence (an illegal consonant cluster) compared to reduced word-forms that start with a phonotactically legal sequence in German.

Measuring participants' ERPs will reveal how listeners process reduced forms and at what point in time difficulties may arise when hearing a reduced form. Floccia et al. (2006) examined the time course of adaptation to a regional accent within the course of an experiment (across blocks). Using ERPs, we can look at the time course of adaptation within trials. Listeners may, for instance, only be disturbed by reduced forms in an early time window (and show an N200 effect), but be able to access the meaning of the word nevertheless (and thus show no N400 effect). This might be the case when processing unfamiliar, but naturally occurring *ge*-reductions. The *ver*-reductions, where more segments are omitted, may pose greater difficulties for processing and thus elicit both N200 and N400 effects. Furthermore, based on the findings by Spinelli and Gros-Balthazard (2007), *ge*- and *ver*-reductions that result in illegal phonotactic sequences may elicit early effects (in the N200 time window), but then later on (in the N400 time window) show patterns like canonical words. These effects may potentially be stronger for *ver*-reductions, as they occur rarely.

Depending on experience with a certain type of reduction, processing difficulties may just be temporary or indeed longer-lasting. As the Southern Germans are very familiar with *ge*-reductions, they should not show any difficulties in processing them and thus not exhibit an N200 or N400 effect. We expect that the Northern Germans will have most difficulties with *ge*-reductions, as they are least familiar with them. Northern German participants should therefore show an N200 and/or N400 effect for reduced *ge*-words. The Northern Germans that were living in the South (the migrant group) are more experienced with *ge*-reductions than the non-migrant Northern German group. They might even have adapted totally to *ge*-reductions and process them as the Southern Germans do. Therefore, the migrants might show a pattern similar to the Southern Germans or at least a pattern in between the Northern and the Southern Germans (e.g., an N200, but no N400 effect).

In contrast to the *ge*-reductions, the *ver*-reductions are equally uncommon in all parts of Germany. ERP responses to reduced *ver*-prefixes may therefore be similar across groups. As neither group is very familiar with them, they should pose more difficulties for processing and thus elicit stronger effects than the *ge*-reductions and possibly both N200 and N400 effects. It is, however, also conceivable that listener groups who are used to dealing with prefix reductions (i.e., the Southerners and maybe also the migrants) can generalize their knowledge to this relatively rare reduction type and may show only temporary difficulties with *ver*-reductions. That is, we may find a group difference not only for the *ge*- but also for the *ver*-reductions.

As the expert group with respect to *ge*-reductions, the Southern Germans should not need to adapt. We therefore do not expect to find short-term adaptation effects for this group. For the Northern German group and maybe also the migrant group, we might observe that the N200s and/or N400s get less negative in the second half of the experiment compared to the first half. All three groups might show less negative components in the second half for the reduced *ver*-words as well.

Corpus studies were conducted to check whether our assumptions about the frequency of occurrence of *ge*- and *ver*-reductions in Northern and Southern Germany were correct. While the Kiel Corpus of Spontaneous Speech confirmed that the Northerners neither reduced the *ge*-prefix to [g]/[k] nor the *ver*-prefix to [f] (the usual pronunciation was [fɐ], which is also the assumed canonical pronunciation in the Kiel Corpus), no such data was available for the Southern German accents. Therefore, we analyzed the audio track of three short movies of the series “The world in Swabian” (Dodokay, 2010a, 2010b, 2010c) transmitted in South Western German television between 2009 and 2011. Swabian is the dialect spoken in South Western Germany (i.e., also in Tübingen, where most of the EEG data were collected). In the series “The world in Swabian”, excerpts of German news, talk shows and the like are made fun of and dubbed into Swabian. Everything is spoken by one male native speaker of Swabian (Kuhn, 2005). As the series is a conscious production of the Swabian dialect, some pronunciations may be exaggerated. All instances of *ge*-past participles and *ver*-verb forms (i.e., past participles and other verb forms, like infinitives) occurring in the three selected movies were analyzed. Out of 19 *ge*-participles, all had the prefix-schwa deleted. If the stem started with a stop, the entire prefix was omitted (e.g., *gekauft* ‘bought’ pronounced as *kauft*). Out of 14 *ver*-verb forms, all prefixes were realized containing two segments, [f] and a vocalic segment. That is, while the Southerners are very likely to omit the vocalic part of the *ge*-prefix, they are far less likely to do the same with the *ver*-prefix. However, when comparing the formant values of the vocalic element in *ver*- with those of a schwa and the canonical [ɐ], some utterances were more schwa-like. That is, the vowel in *ver*- is lenited but not deleted as in *ge*-. This suggests that only Southerners but not Northerners have extensive experience with the omission of the schwa in *ge*- and neither group is at all familiar with utterances of *ver*- in which the vowel is completely deleted.

Method

Participants

One hundred and sixty-six native speakers of German took part in this experiment. According to their origin, they were categorized as either Northern or Southern Germans. The notion "Northern" was taken in a very broad sense including also the Eastern and Western parts of Germany, but excluding the states of Baden-Württemberg and Bavaria. Only participants who had grown up in Austria or in the German states of Baden-Württemberg and Bavaria were regarded as Southern German speakers. Additionally, we distinguished Northern Germans according to their contact with the Southern German variant. Forty-two participants (20 males), categorized as Northern Germans, were tested at the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands. They formed the non-migrant Northern German group, with little contact with the Southern German variant. A second group of Northern Germans living in the South (41 participants, 13 males), called the migrant Northern German group, constituted a group of Northern German supposedly familiar with the Southern variant. The Southern German group (42 participants, 20 males), the migrant group and the control group (41 participants, 20 males) were all tested at the Eberhard Karls

University of Tübingen, Germany. Tübingen is situated in the Southwestern part of Germany, in the state of Baden-Württemberg.

According to a questionnaire that was given to all participants after the ERP-experiment, participants in the Southern group were very familiar with *ge*-reductions (93% indicated that they hear *ge*-reductions in their environment; 57% indicated that they produce *ge*-reductions) and also somewhat familiar with *ver*-reductions (26% claimed to have been exposed to such reductions, while 5% claimed to produce them). Participants in the migrant group were more familiar with *ge*-reductions than the non-migrant Northern group (migrants: 63% exposure; 0% usage; non-migrant Northerners: 19% exposure; 2% usage), but did not have more experience with *ver*-reductions (migrants: 17% exposure; 0% usage; non-migrant Northerners: 14% exposure; 2% usage). None of the participants reported any hearing disorders, all had normal or corrected-to-normal vision and were right-handed. None had any neurological impairment or had experienced any neurological trauma. The data of additional 20 participants (nine in the control group, two in the non-migrant Northern German group, two in the migrant Northern German group, and seven in the Southern German group) were not analyzed due to strong alpha waves or very noisy data, leading to the rejection of more than 50% of their data. Participants were paid for taking part.

Design

We targeted ERP components that are sensitive to the differences between words and nonwords in spoken language comprehension. Participants listened to 400 sentences which contained various types of past participles. In sentences containing *be*-participles (i.e., past participles starting with a derivational *be*-prefix, e.g., *besiegt* 'defeated'), the participles were always unreduced. These sentences formed the baseline condition for possible N400 effects. Nonwords were created which were phonologically related to the *be*-participles (see Materials for more details). ERP responses to the sentences containing nonwords provided another comparison condition. Sentences ending in reduced *ge*- and reduced *ver*-participles were expected to elicit brain responses somewhere along the continuum between responses to canonical *be*-participles and responses to nonwords. Ideally, the processing of the reduced *ge*- and *ver*-participles would be compared (within participants) to the processing of unreduced participles containing the same prefix. However, when a listener encounters both reduced and unreduced forms of a given prefix, this very likely increases the risk that an adaptation effect over the course of the experiment will not be observed. Therefore, the processing of unreduced participles containing a different prefix (*be*-) was chosen as the baseline.

In order to determine a possible long-term adaptation effect to the reduced prefixes, three groups of participants were tested: Northern Germans, Southern Germans and Northern Germans living in the South. These three groups listened to exactly the same materials: sentences containing canonical *be*-participles, nonwords, reduced *ge*-participles and reduced *ver*-participles. A fourth group was tested who did not hear any reductions. Participants in this control group listened to the same sentences ending in canonical *be*-participles, nonwords, canonical *ge*-participles and canonical *ver*-participles.

Materials

The experimental items consisted of five sets of 80 sentences each: one set of sentences containing *be*-participles, one containing nonwords instead of past participles, one containing *ge*-participles, one containing *ver*-participles and one set of filler sentences, which did not contain any past participles. Apart from the 80 filler sentences, all sentences ended either in a past participle or in an auxiliary verb preceded by a past participle (see appendix for a full list of the past participles and the nonwords used in the experiment). The *ge*-participles were past participles formed with the inflectional *ge*-prefix (e.g., *Für diese Prüfung habe ich echt stundenlang <u>ge</u>lernt.* 'I have learned for hours for this exam'). Verbs that already have a derivational prefix, like *ver*- or *be*- (e.g., *verwirren* 'to confuse' or *besiegen* 'to defeat'), do not need the inflectional *ge*-prefix to form the past participle (e.g., *Mit seinen Erklärungen hat mich unser Mathelehrer nur noch mehr <u>ver</u>wirrt.* 'Our maths teacher has only confused me more with his explanations.' or *Nach monatelanger Chemotherapie hat Anjas Mutter den Krebs endlich <u>be</u>siegt.* 'Anja's mother has finally defeated cancer after months of chemotherapy'). Sentences were constructed in such a way that the past participles were predictable (see cloze test below).

Half of the 80 *ge*-participles resulted in illegal phonotactic sequences when they were reduced. That is, omitting the schwa in [gəlɛʀnt] *gelernt* 'learned' leads to a phonotactically legal German sequence, whereas omitting the schwa in [gəʃvɔmən] *geschwommen* 'swum' leads to a phonotactically illegal sequence in German. The same holds for omitting the vocalic part of the prefix *ver*- [fɛɐ]. The onset of reduced [floːʀən] for *verloren* 'lost' is phonotactically legal, while the onset of reduced [fvɪʀt] *verwirrt* 'confused' is not. For the *ver*-words, it was only possible to create a legal set of 33 items. In order to balance the design, the legal set was filled up with seven *ver*-words that actually led to an illegal sequence when reduced (e.g., *verhaftet* 'arrested' > *fhaftet*). The stem of these seven *ver*-words started always with /h/, a sound that is perceptually less salient, especially in the context of voiceless obstruents, which induce some (phonetic) aspiration themselves. The consonant cluster /fh/ did not occur in the illegal set. Note that by treating some illegal *ver*-words as legal in the analyses, we are making it harder to observe a difference in processing between reductions leading to legal sequences and those leading to illegal sequences. The *be*-participles contained consonants in third position that would result in phonotactically legal sequences if the schwa of the prefix were to be omitted. The legal onset clusters /ks/ (*ge*-words), /fh/ (*ver*-words), and /ps/ (*be*-words) are quite infrequent. Overall, 20% of the legal words started with such a infrequent legal cluster.

Eighty nonwords modeling schwa-deletions were constructed on the basis of the *be*-participles. The schwa following the initial /b/ was removed. If a voiceless consonant was in third position, the initial /b/ was devoiced. The vowel of the stem of the base participle was also changed. For example, on the basis of [bəziːkt] *besiegt* 'defeated' the nonword [psuːkt] *psugt* was created. (Note that Southern German speakers tend to devoice /z/ in onset position.) The sentence-frames for the *be*-participles and their corresponding nonwords were identical (e.g., *Nach monatelanger Chemotherapie hat Anjas Mutter den Krebs endlich besiegt/psugt*). Apart from the *ge*-targets and *ver*-targets, the sentences did not contain any other *ge*- or *ver*-prefixes.

Cloze test

A cloze test was run to determine the degree of predictability of the sentence-final past participles. Nineteen native German speakers, none of whom participated in the main experiment, were paid to take part. They were seated in a sound-proof booth in front of a computer screen. The 240 sentences containing *be*-, *ge*- and *ver*-participles were presented in a randomized order with the participle replaced by a gap. Participants were instructed to complete these sentences with one word. They were asked to type in at least one answer, but had the possibility to give up to seven answers. After typing in their answer(s), participants saw the same sentence again, this time complete with the past participle used in the ERP experiment. They were asked to rate how well the proposed past participle completed the sentence context on a scale from 1 ("Word does not fit at all") to 7 ("Word fits perfectly"). The cloze test was self-paced; it took participants 60 to 90 minutes.

We analyzed the percentages of mentioned target words in the sentence completion task, the percentages of mentioned words that started with the same prefix as the target word, and the mean ratings for the targets in the rating task. The *ge*-participles that were used in the main EEG experiment were mentioned in 86% of the cases, the *ver*-participles were mentioned in 57% and the *be*-participles in 43% of the cases. A similar pattern was found considering all responses (i.e., the target and any other words): If a *ge*-participle was the target, on average 85% of participants' responses were words starting with a *ge*-prefix. If a *ver*-participle was the target, on average 62% of participants' responses were words starting with a *ver*-prefix. If a *be*-participle was the target, on average 46% of participant's responses were words starting with a *be*-prefix. This seems to reflect a frequency effect as participles containing the derivational prefixes *ver*- and *be*- were mentioned less frequently than participles built with the inflectional *ge*-prefix. On average, the *ge*-past participles were most frequent (their form frequency was 2.6 per million, based on the SUBTLEX-DE corpus by Brysbaert et al. (2011)), followed by the *ver*-past participles (1.9 per million) and the *be*-past participles (1.5 per million). The difference in percentages for mentioning the *ver*- and *be*-targets can probably also be explained by the higher constraints on the *be*-words. The *be*-past participles formed the basis for the nonwords. The nonwords were designed to model a schwa-reduction of the *be*-prefix but result in a phonotactically legal sequence in German. Therefore, the choice for the first consonant of the stem was constrained (to /l/, /r/, /h/ and /s/) and it had to be followed by a vowel (e.g., the nonword based on *beleidigt* 'insulted' was *blodigt*). Although participants did not come up with our solutions in all cases, they rated them very highly across all three types of past participles: On a scale from 1 to 7, with higher ratings meaning better fits, participants rated the *ge*-participles with 6.9, the *ver*-participles with 6.5 and the *be*-participles with 6.3 on average.

Stimulus construction

The experimental sentences were recorded digitally with a sampling rate of 44.1 kHz by a female native speaker of German from Austria sitting in a sound-attenuating booth. The speaker was chosen for her southern origin so that she would be able to produce the *ge*-reductions naturally. She was instructed to pronounce the sentences canonically. Each sentence was recorded twice, once ending in an unreduced (*be*-, *ge*-, or *ver*-) word and once ending in a nonword or a reduced (*ge*- or *ver*-) word. This procedure was chosen to avoid cross-splicing of the stimulus material. For the reduced versions and the nonwords, the targets were transcribed orthographically (e.g., she got to read *Mit seinen Erklärungen hat* *mich unser Mathelehrer nur noch mehr fwirrt* with *fwirrt* replacing *verwirrt* of the canonical sentence) and she was told to read them as naturally as possible. The sentences were printed out on several pages, and per page, first the sentences containing canonical words were recorded followed by the same sentences containing reduced past participles or nonwords.

Procedure

The Northern German group was tested in Nijmegen, The Netherlands, while the other three groups were tested in Tübingen, Germany. The procedure was the same for all participants. They were seated in a comfortable chair in front of a computer screen and told that they would hear 400 sentences. They were asked to listen carefully and to answer questions about the content of some sentences by pressing one of two possible response buttons, when a question and the two response options appeared on the screen. The questions concerned the content of 40 filler sentences and were included to ensure participants' attention. Participants were instructed to blink or move in any other way as little as possible when a fixation cross was displayed on the screen.

During each trial, the trial number was shown first for 500 ms. Then, a fixation cross appeared, while the spoken sentence was being played. If the current trial was a filler trial with a question to answer, the question and two response possibilities appeared on the screen next. The question stayed on the screen until the participant pressed one of the two response buttons. Finally, an instruction to blink appeared on the screen for 2000 ms. Inspection of the data of the first eight participants (all members of the Northern German group) revealed that participants already prepared to blink before the auditory sentence finished. Therefore, the run of a trial was changed slightly: The fixation cross was now displayed for another 800 ms after the end of the spoken sentence. Participants could take a longer break after every 50th trial. The experimenter checked on participants during the breaks, offered something to drink and chatted with them in order to keep them alert. When participants had completed half of the experiment, they were offered a chocolate bar. After the experiment, participants were asked to fill in a questionnaire. The order of the 400 sentences was randomized for each participant individually. After completing the ERP-experiment, participants were asked to fill in a questionnaire. We were interested in where they were born and raised, whether and how long they had lived in Southern Germany or Austria, whether they spoke a dialect, whether they were familiar with *ge*- and *ver*-reductions, whether they would produce them themselves, and how often *ge*- and *ver*-reductions were produced in their environment. An experimental session took approximately two hours, while the actual ERP-experiment took about 60-65 minutes.

EEG-recording and analysis

In Nijmegen, the EEG was recorded with BrainVision Recorder (version 1.03.0004) from 36 active Ag/AgCl electrodes, all of which but one were mounted in a cap (actiCap). One separate electrode was placed on the infraorbital ridge below the left eye and was used for monitoring blinks. Horizontal eye movements were monitored through two electrodes in the cap (F9, F10). The ground electrode was situated at the forehead. Electrodes were referenced to the left mastoid electrode. Electrode impedance was kept below 20 kΩ, which is a sufficiently low impedance when using active electrodes. EEG and EOG recordings were amplified through BrainAmp DC amplifiers using a bandpass filter of 0 - 100 Hz, digitized on-line with a sampling frequency of 500 Hz, and stored for off-line analysis. The EEG data from Nijmegen were later re-sampled to a frequency of 512 Hz to be comparable with the EEG data collected in Tübingen.

In Tübingen, the EEG was recorded with BioSemi ActiView (version 6.05) from 69 active Ag/AgCl electrodes, of which 64 were mounted in a cap. Blinks were monitored through an electrode placed on the infraorbital ridge below the left eye. Horizontal eye movements were monitored through two separate electrodes (LEOG, REOG) placed approximately at each outer canthus. Another two electrodes were placed at the left and right mastoids. In the BioSemi system, the conventional ground electrode is replaced by two separate electrodes: the Common Mode Sense (CMS) active electrode and the Driven Right Leg (DRL) passive electrode. The former was located in the cap between the electrodes PO3 and POz, the latter was located between the electrodes POz and PO4. The EEG data saved to file were unreferenced. BioSemi's active electrodes have an output impedance of less than 1Ω. EEG and EOG recordings were amplified through the ActiveTwo AD-box using a bandpass filter of 0.16 - 100 Hz, digitized on-line with a sampling frequency of 512 Hz, and stored for off-line analysis.

The EEG data from both Nijmegen and Tübingen were preprocessed with BrainVision Analyzer (version 2.0.1.3931). All channels were re-referenced to the linked mastoids. Bipolar vertical EOG was computed as the difference between the electrode on the infraorbital ridge of the left eye and electrode Fp1. Bipolar horizontal EOG was computed as the difference between electrodes F9 and F10 for the Nijmegen data and as the difference between the LEOG and REOG electrodes for the Tübingen data. The data were corrected for eye-movement artifacts using Independent Component Analysis (ICA). The EEG was band filtered between 0.5 Hz (12 dB/oct) and 25 Hz (24 dB/oct). The signals were segmented into epochs of 1000 ms (with a 200 ms prestimulus baseline, time-locked to the onset of the [nonword-]past participle). Epochs with an amplitude outside the range of -50 to 50 microvolt were automatically excluded (on average 89.3% of the epochs were kept). Average ERPs were then computed across trials per participant for each condition (*be*-, *ge*-, *ver*- and nonwords) overall, for the experimental halves per condition (e.g., *ge*-words in the first half vs. *ge*-words in the second half of the experiment) and for legal vs. illegal phonotactic sequences for *ge*- and *ver*-words (e.g., legal *ge*-words vs. illegal *ge*-words). These averages were used to create different difference waves for each of the analyses (see Table 1).

Table 1. Difference wave forms that were created for the different types of analysis.

Effects	Difference waves		
	nonwords	*ge*-words	*ver*-words
Long-term adaptation	nonwords - *be*-words	*ge*-words - *be*-words	*ver*-words - *be*-words
Short-term adaptation	1st half - 2nd half	1st half - 2nd half	1st half - 2nd half
Phonotactic legality		legal – illegal	legal - illegal

Results

Accuracy data

Participants in all groups answered 76-79% of the filler questions correctly.

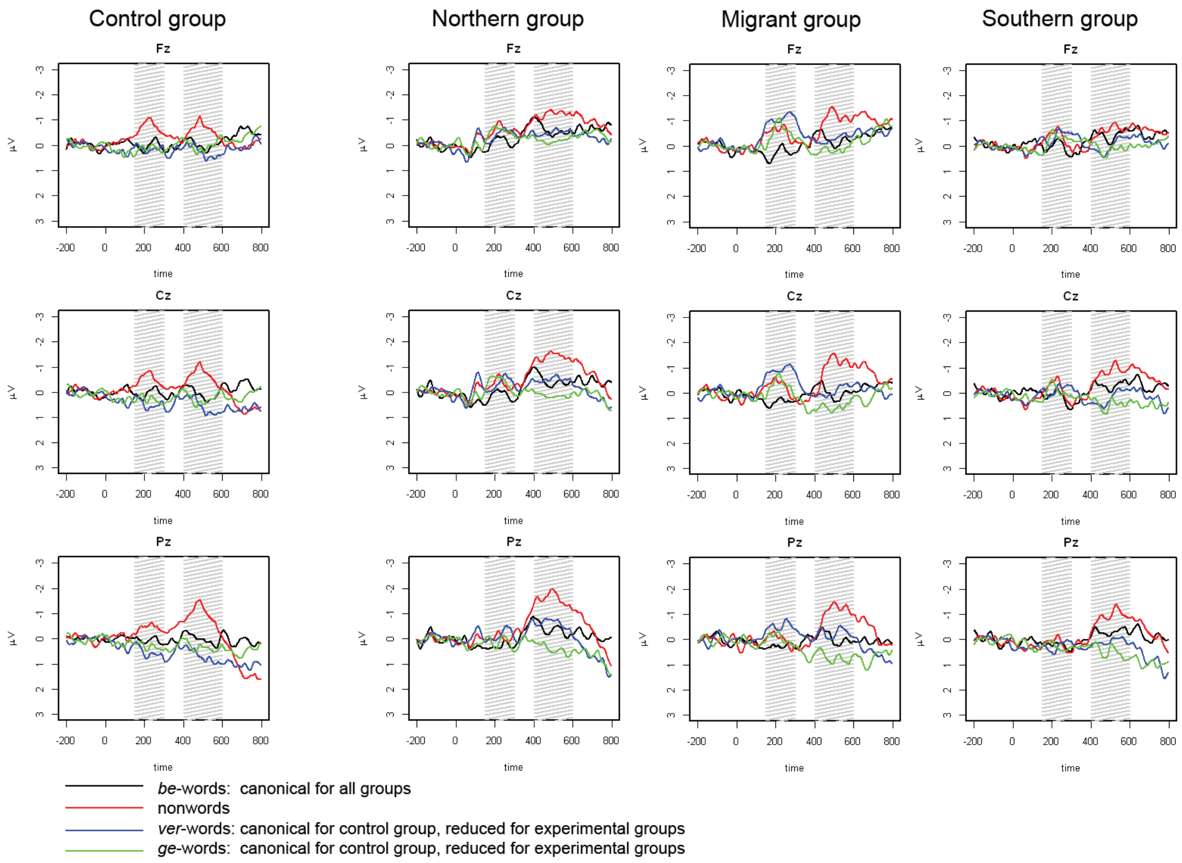
ERP data

Analyses were conducted using repeated measures ANOVA for the three midline electrodes Fz, Cz and Pz. Where necessary, *p*-values are based on Greenhouse-Geisser corrections. In these cases, Greenhouse-Geisser ε is also reported.

Figure 1 shows the ERPs to all four word types (*be*-words, nonwords, *ge*-words, *ver*-words) by all four groups (control group, non-migrant Northern German group, migrant Northern German group, Southern German group) for the three midline channels (Fz, Cz, Pz). For the control group, who heard only canonical forms and nonwords, the nonwords gave rise to two negative deflections, one around 200 ms and another one starting around 400 ms, in comparison to the other conditions. These deflections with respectively frontal and parietal maxima seem to reflect the well-known components N200 and N400. We therefore analyzed two time windows, the first one ranging from 150-300 ms and the second ranging from 400-600 ms (Van den Brink, et al., 2001; see the shaded areas in Figure 1). To facilitate the comparison of the groups, we analyzed difference waves between the control condition (canonical *be*-words) and the three experimental conditions.

Examining long-term and short-term adaptation effects, we did not compare the processing of the reduced *ge*-words to the processing of the reduced *ver*-words within groups, because the unreduced *ge*-words and the unreduced *ver*-words differed in frequency (the *ge*-words were more frequent than the *ver*-words) and in cloze probability (the *ge*-words were more predictable than the *ver*-words). Instead we focused on the comparison of a given condition (*ge*-words or *ver*-words) across different versions (reduced or canonical variants) and across different listener groups.

Figure 1. ERPs to all four word types (*be*-words, nonwords, *ge*-words, *ver*-words) by all four groups for the three midline channels. The *ge*-words and *ver*-words were not reduced for the control group, while they were for the three other groups.



N200 time window. The average difference waves in the early time window were subjected to repeated measures ANOVA with the factors Channel (Fz, Cz, Pz), Condition (nonwords, *ge*-words, *ver*-words), and Group (control group, non-migrant Northern German group, migrant Northern German group, Southern German group). The three-way interaction was not significant ($F(12, 648) = 1.53$, $\varepsilon = 0.81$, $p > .1$). However, there was a significant interaction between Group and Condition ($F(6, 324) = 5.43$, $\varepsilon = 0.96$, $p < .001$). The relevant means and their standard errors are displayed in Figure 2. To investigate the interaction, we examined the effect of Group for each Condition. ANOVAs with Group as independent variable revealed the following patterns.

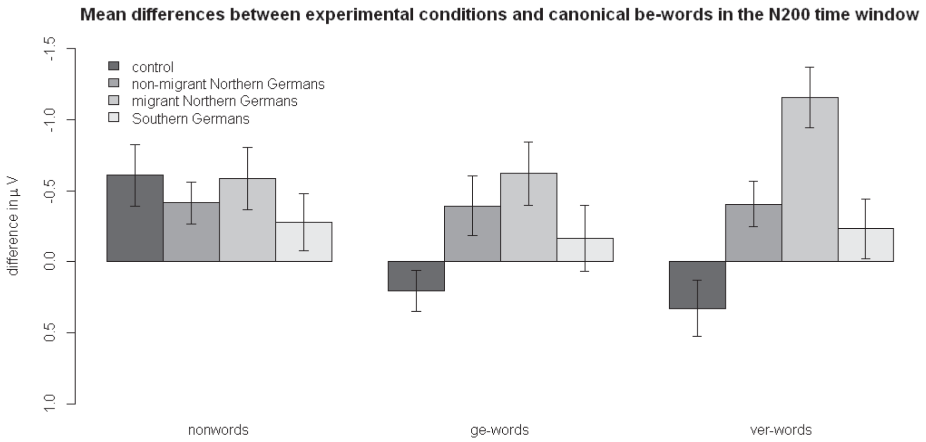
First, there was no group difference for the processing of the nonwords ($F(3, 494) = 1.66$, $p > .1$). Given the absence of a group difference, we tested whether the overall difference of -0.47 µV was significantly different from zero. This was the case ($t(165) = -4.78$, $p < .001$). This means that the nonwords, which were presented in identical forms to all four groups, generated a similar N200 effect in all groups.

Second, for the *ge*-words, the ANOVA with Group as independent variable revealed a significant effect ($F(3, 494) = 7.57$, $p < .001$). Post-hoc Tukey HSD tests showed significant differences between the control group and the two Northern German groups (non-migrant group: $p < .01$, migrant group: $p < .001$), with a larger N200 for the two Northern German groups. Note that this difference was driven by different stimuli, as the control group heard canonical *ge*-words while the two Northern German groups heard reduced *ge*-words. The Southern Germans, despite hearing reduced forms as well, did not differ from the control group (p > .1). That is, as predicted, despite hearing a reduction, they did not show an N200 effect. Furthermore, there was a marginal difference (p = .057) between the Southern Germans and the migrant Northern Germans, with a larger N200 in the migrant Northern German group.

Third, for the *ver*-words, the ANOVA with Group as independent variable also showed a significant difference between groups ($F(3, 494) = 24.83$, $p < .001$). Post-hoc comparisons with Tukey HSD tests revealed that the ERP data of the control group differed significantly from those of the three other groups (non-migrant Northern Germans: $p < .001$, migrant Northern Germans: $p < .001$, Southern Germans: $p < .01$). That is, the reduced words presented to the latter three groups elicited more negative difference waves than the unreduced forms presented to the control group. There was also a significant difference between the migrant Northern German group on the one hand and the non-migrant group ($p <$

.001) and the Southern German group ($p < .001$) on the other, with the largest N200 for the migrant Northern German group.

Figure 2. Mean differences and SEs between experimental conditions (nonwords, *ge*-words, *ver*-words) and canonical *be*-words in the N200 time window averaged over the three midline channels.

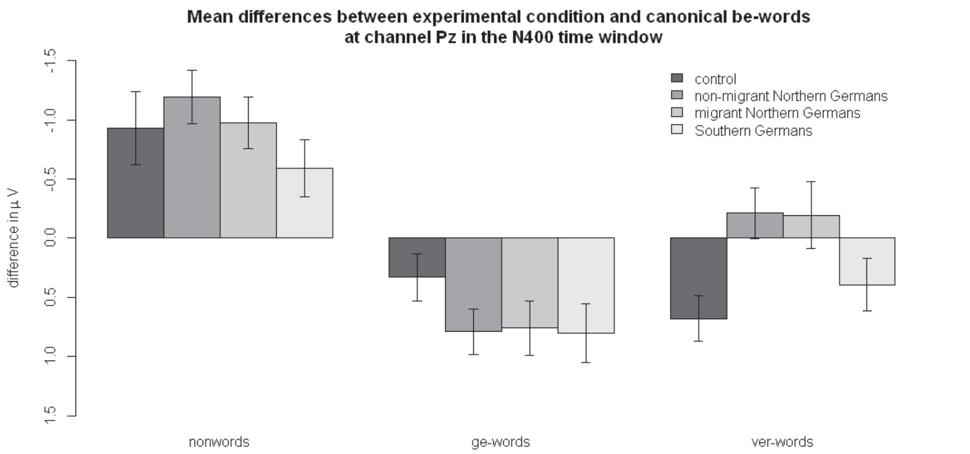


N400 time window. The average difference waves in the late time window were also subjected to repeated measures ANOVA with the factors Channel (Fz, Cz, Pz), Condition (nonwords, *ge*-words, *ver*-words), and Group (control group, non-migrant Northern German group, migrant Northern German group, Southern German group). In the N400 time window, the three-way interaction was significant ($F(12, 648) = 1.91$, $\varepsilon = 0.85$, $p < .05$). We further explored this interaction by running ANOVAs with Condition and Group as independent variables. There was a main effect of Condition at channels Fz ($F(2, 324) = 33.48$, $\varepsilon = 0.94$, $p < .001$) and Cz ($F(2, 324) = 51.41$, $\varepsilon = 0.97$, $p < .001$). At channel Pz, the two-way interaction between Condition and Group was significant ($F(6, 324) = 3.39$, $\varepsilon = 0.95$, $p < .01$). As the effects in the N400 time window are most prominent at channel Pz (see Figure 1), we focus the further analysis on this channel.

The two-way interaction at Pz was examined by running an ANOVA with Group as independent variable for each level of Condition. The relevant means and their standard errors are displayed in Figure 3. First, for the nonwords, we did not find a significant difference between groups ($F(3, 162) = 0.99$, $p > .1$). A further test whether the overall difference of -0.92 µV was different from zero was significant ($t(165) = -7.37$, $p < .001$). That is, as predicted, all groups showed a similar N400 effect for the nonwords at Pz. Note that this effect emerged although the carrier sentences for *be*-words and nonwords were repeated.

Second, for the *ge*-words, there was no significant differences between groups either ($F(3, 162) = 1.07$, $p > .1$). Testing the overall difference of 0.67 µV against zero showed a significant effect ($t(165) = 6.14$, $p < .001$). The positive overall difference probably reflects that the *ge*-words were somewhat more predictable than the *be*-words. Nevertheless, the ERP data at channel Pz suggest that all groups processed the *ge*-words similarly. There was no difference in the processing of the *ge*-words for the control group (who heard canonical versions) compared to the three other groups (who heard reduced versions). None of the experimental groups showed more negative difference waves than the control group in the N400 time window, whereas the two Northern German groups showed exactly that in the N200 time window.

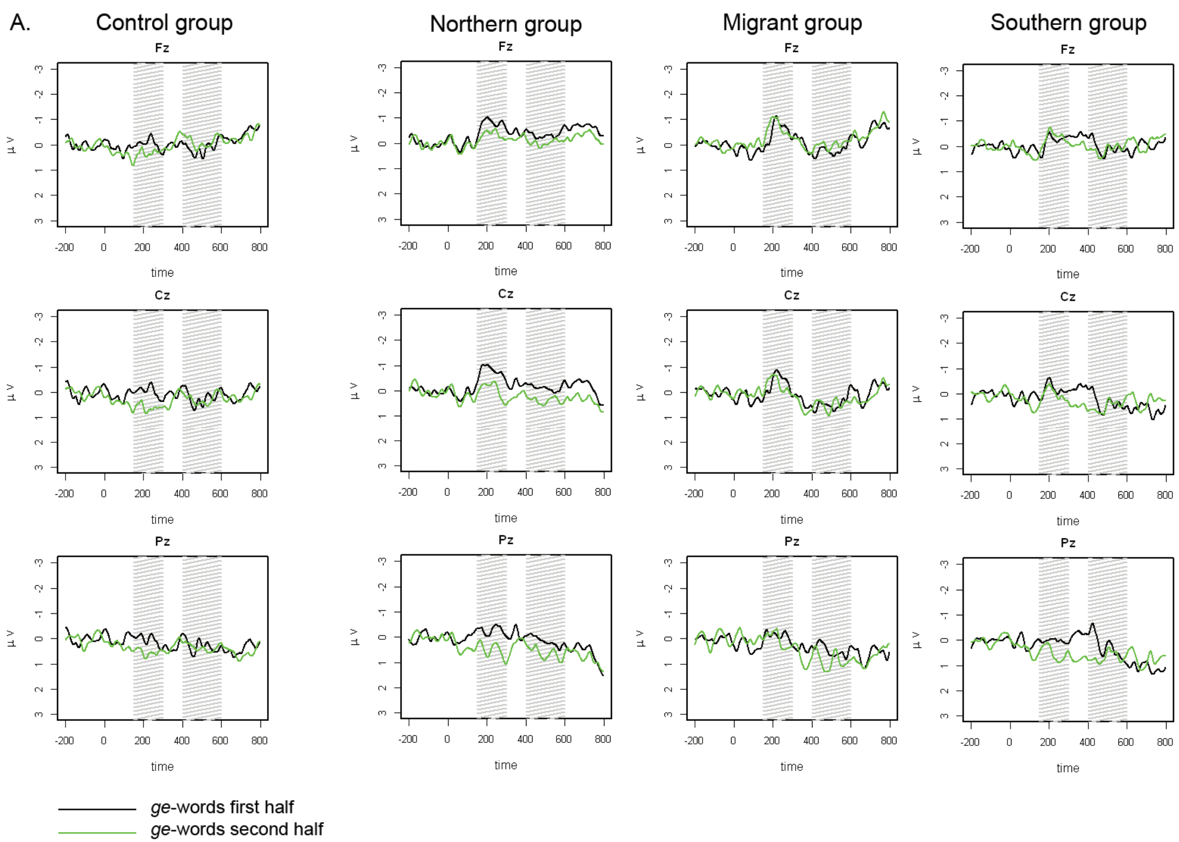
Figure 3. Mean differences and SEs between experimental conditions (nonwords, *ge*-words, *ver*-words) and canonical *be*-words in the N400 time window at channel Pz.

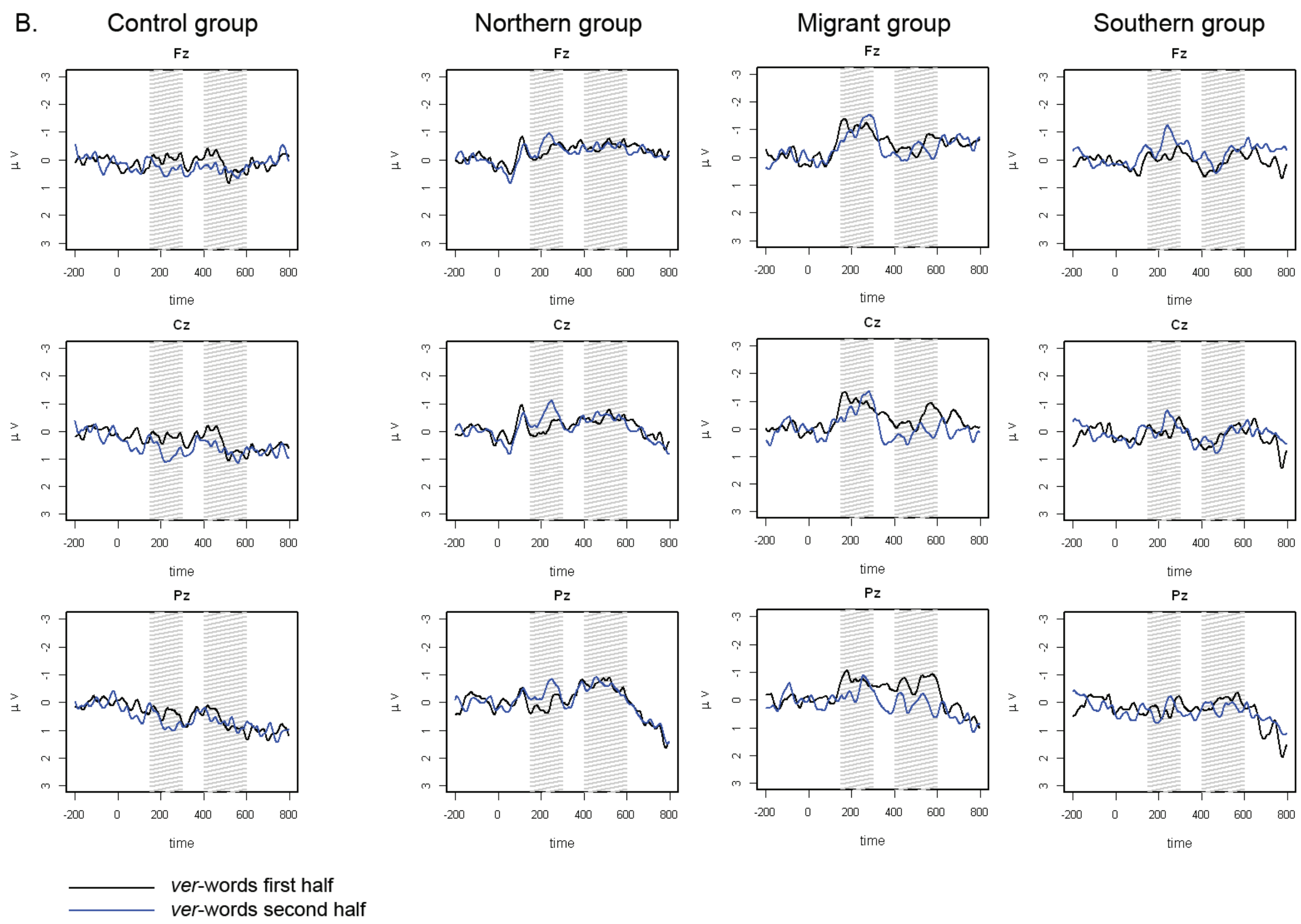


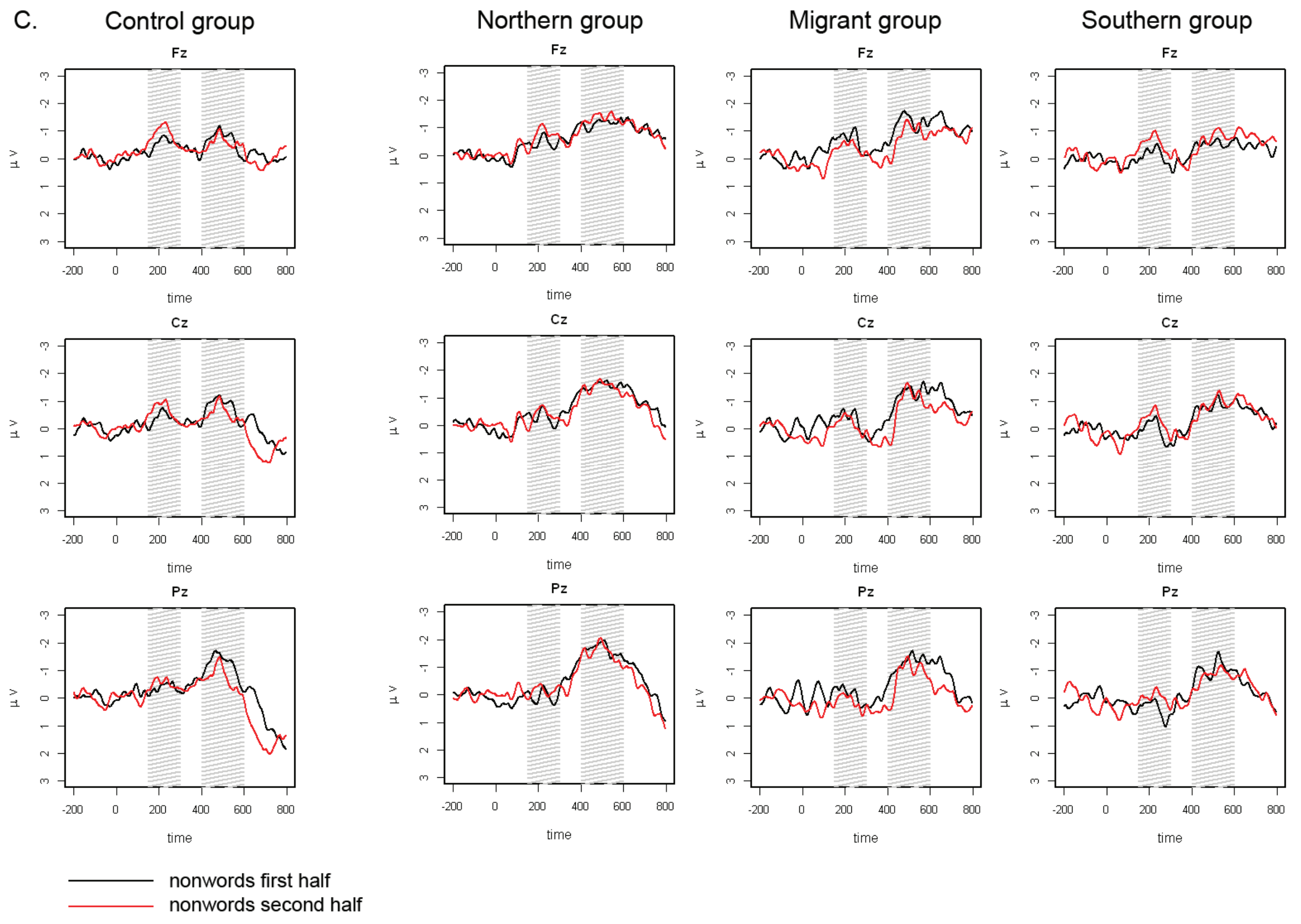
Third, for the *ver*-words, we found a significant difference between groups ($F(3, 162) = 3.68$, $p < .05$). Post-hoc tests with Tukey HSD revealed that the ERP data of the two Northern German groups differed significantly from those of the control group (non-migrant Northern Germans: $p < .05$, migrant Northern Germans: $p < .05$) with more negative difference waves in the two Northern German groups. This difference is driven by different stimuli, as the control group heard canonical *ver*-words while the two Northern German groups heard reduced *ver*-words. The Southern German group, despite hearing reduced forms as well, did not differ from the control group ($p > .1$), but also did not differ significantly from the two Northern German groups.

Short-term adaptation effects. Figure 4 shows the ERPs for the nonwords, *ge*-words and *ver*-words split into experimental halves by all four groups (control group, non-migrant Northern German group, migrant Northern German group, Southern German group) for the three midline channels (Fz, Cz, Pz). If listeners adjusted to the reduced stimuli in the course of the experiment, we should see a smaller N200 and/or N400 in the second half of the experiment compared to the first half of the experiment. In order to establish that any changes were really adjustments and not simply effects of time (i.e., participants getting tired over the course of the experiment), the difference waves between the first and second half of the reduced conditions were compared to the difference waves of the nonwords. This comparison allows us to isolate adjustment effects from simple effects of time (such as fatigue), which should affect nonwords and the prefixed words to the same degree. Moreover, the results show adjustment effects to reductions only if the change over the course of the experiment is larger in the experimental groups than in the control groups, since only the former heard reductions. According to this criterion, we found no short-term adaptation effects to reduction.

Figure 4. ERPs for the *ge*-words (panel A), the *ver*-words (panel B) and the nonwords (panel C) split into experimental halves by all four groups for the three midline channels. The *ge*-words and *ver*-words were not reduced for the control group, while they were for the three other groups.




B.
Control group
Northern group
Migrant group
Southern group
Fz
Cz
Pz
µ V
time
ver-words first half
ver-words second half


C.
Control group
Northern group
Migrant group
Southern group
Fz
Cz
Pz
µ V
time
nonwords first half
nonwords second half

Phonotactic legality effects. Figure 5 displays the ERPs for the *ge*-words split into legality categories (i.e., "legal" and "illegal", according to whether they result in legal or illegal phonotactic sequences in standard German when they are encountered in reduced form) by all four groups (control group, non-migrant Northern German group, migrant Northern German group, Southern German group) for the three midline channels (Fz, Cz, Pz). Figure 6 shows the equivalent for the *ver*-words. Especially in Figure 6, a typical distribution for an N200 and an N400 can be seen, that is, an early negative deflection around 200 ms after target onset with a frontal maximum and a later negative deflection starting around 400 ms after target onset with a parietal maximum. The same time windows as in the previous analyses were used (i.e., the N200 time window ranging from 150-300 ms and the N400 time window ranging from 400-600 ms; see shaded areas in Figures 5 and 6). Difference waves between the legal and the illegal condition were analyzed in order to facilitate the comparison between groups. In both time windows, the average difference waves were subjected to repeated measures ANOVA with the factors Channel (Fz, Cz, Pz), Condition (*ge*-words, *ver*-words), and Group (control group, non-migrant Northern German group, migrant Northern German group, Southern German group).

Figure 5. ERPs for the *ge*-words split into legality categories (i.e., "legal" and "illegal") by all four groups for the three midline channels. The *ge*-words were not reduced for the control group, while they were for the three other groups.

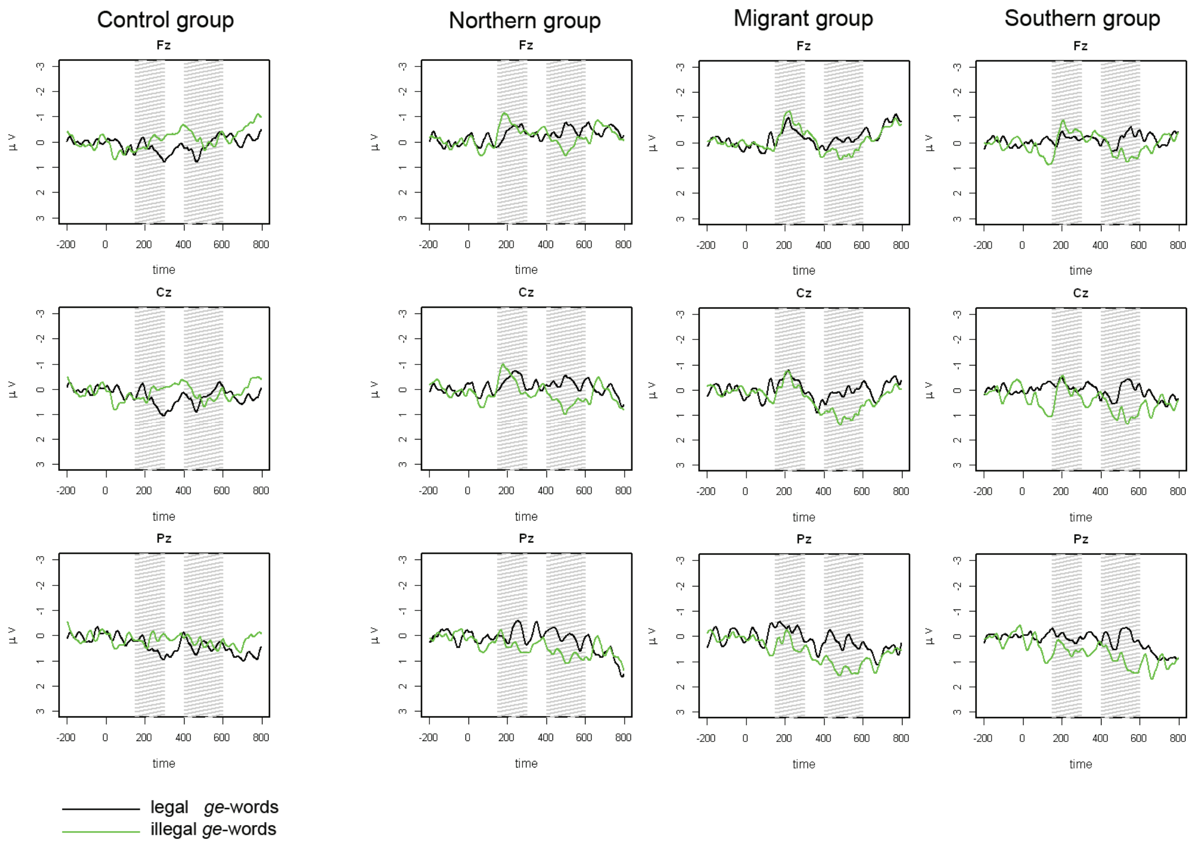
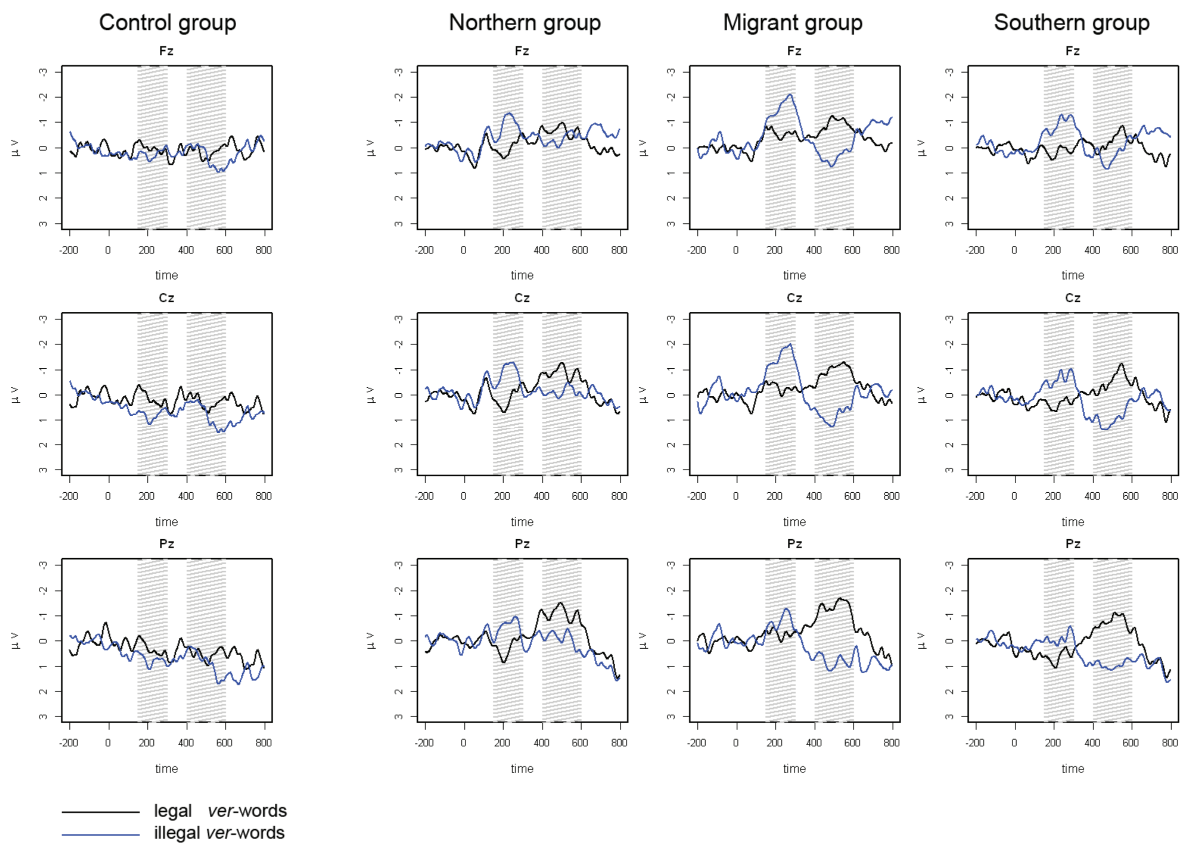
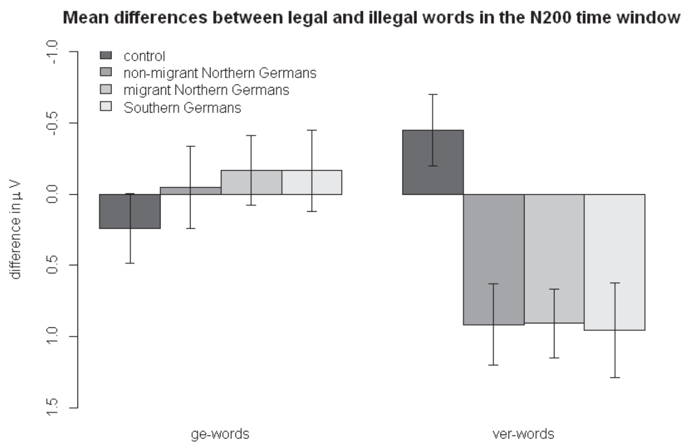


Figure 6. ERPs for the *ver*-words split into legality categories (i.e., "legal" and "illegal") by all four groups for the three midline channels. The *ver*-words were not reduced for the control group, while they were for the three other groups.



N200 time window. In the early time window, the three-way interaction was marginally significant ($F(6, 324) = 1.94$, $\varepsilon = 0.94$, $p = .08$) and the two-way interaction between Group and Condition was clearly significant ($F(3, 162) = 4.69$, $p < .01$). Since the overall analysis showed a Group by Condition interaction for the N200 time window, we focus on this two-way interaction here as well. This two-way interaction was further analyzed by running ANOVAs with Group as independent variable for the *ge*- and *ver*-words separately. The relevant means and their standard errors are displayed in Figure 7. For the *ge*-words, there was no significant difference between groups ($F(3, 494) = 1.33$, $p > .1$). Testing the overall difference of -0.03 µV against zero did not reveal a significant effect ($t(165) = -0.26$, $p > .1$). That is, the groups who heard the *ge*-words in reduced form processed them in a similar way regardless of whether the words resulted in legal or illegal phonotactic sequences and did not differ in their processing from the control group who heard only canonical forms.

Figure 7. Mean differences and SEs between "legal" and "illegal" *ge*-words and *ver*-words in the N200 time window averaged over the three midline channels.



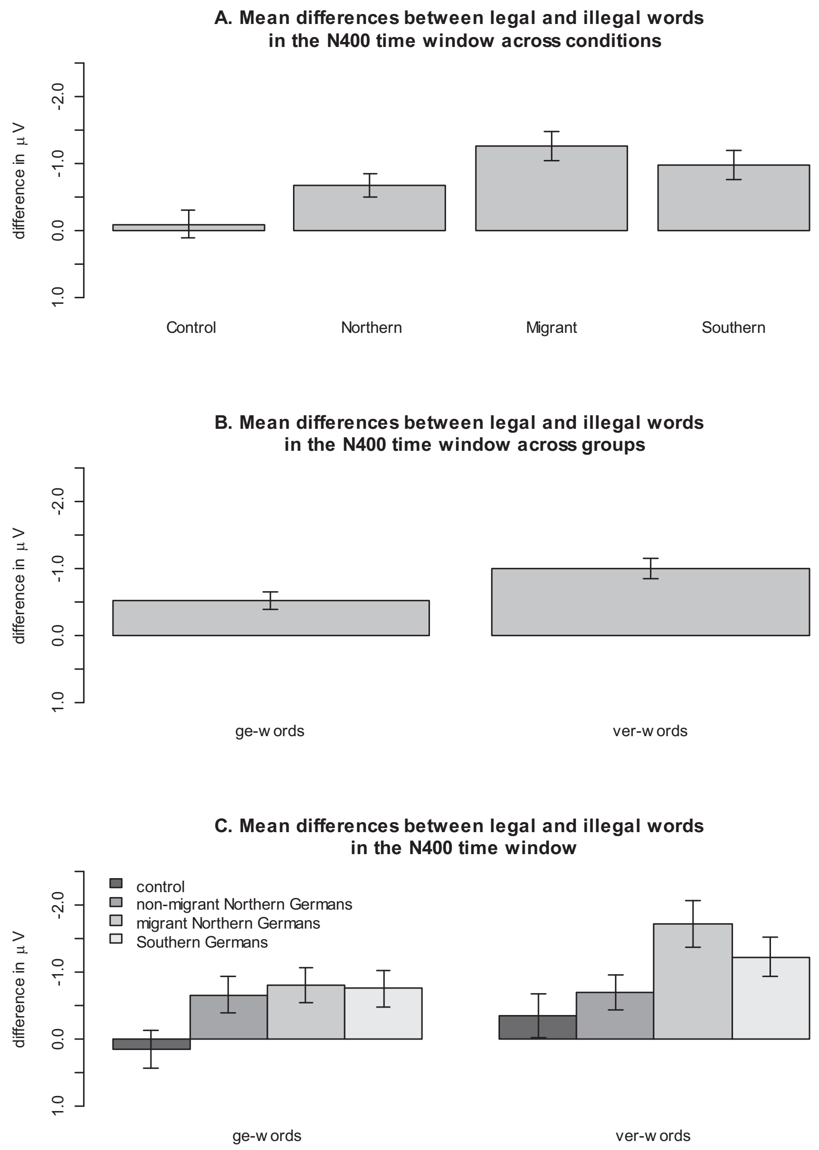
For the *ver*-words, we found a significant difference between groups in processing the legal and illegal items ($F(3, 494) = 15.33$, $p < .001$). Post-hoc comparisons with Tukey HSD showed that the control group differed from the three other groups (all *p*s < .001). Note that this group difference is driven by different stimuli, as the control group only heard canonical *ver*-words while the three other groups all heard reduced *ver*-words. The non-migrant Northern Germans, the migrant Northern Germans and the Southern Germans did not differ in their processing of the (il-)legal *ver*-words, but all groups showed a larger N200 for illegal than for legal clusters.

N400 time window. In the late time window, neither the three-way interaction (F(6, 324) = 0.95, ε = 0.88, $p > .1$) nor the two-way interaction between Group and Condition was significant (F(3, 162) = 0.77, $p > .1$). The main effects of Group (F(3, 162) = 5.81, $p < .001$) and of Condition (F(1, 162) = 5.68, $p < .05$), however, were significant. The relevant means and their standard errors are displayed in Figure 8 (Panels A and B), with additional panels for the marginal means for Group and Condition (Panel C). Further analyses of the main effect of Group with Post-hoc Tukey HSD tests showed that the control group differed from the other three groups (non-migrant Northern German group: $p < .01$, migrant group Northern German group: $p < .001$, Southern German group: $p < .001$). This group difference is due to different stimuli, as the control group heard both the *ge*-words and the *ver*-words in canonical form, whereas the other three groups heard them in reduced form. Furthermore, the two Northern German groups differed in the late time window in their processing of the (il-)legal *ge*- and *ver*-words ($p < .01$) insofar as the migrant group showed a greater difference in processing between the legal and the illegal words than did the non-migrant group.

As for the main effect of Condition, all groups differed in processing the (il-)legal *ge*-words and the (il-)legal *ver*-words ($p < .001$) insofar as the difference in processing between the legal and the illegal words was greater for the *ver*-words than for the *ge*-words.

As Figures 5, 6, and 8 show, the N400 is reduced and even becomes positive when the reduced cluster is illegal (for both *ge*- and *ver*-reductions). Note that, for the *ver*-words, the effect of phonotactic legality changes in direction. In the early time-window, illegal clusters gave rise to a larger N200 than legal clusters, indicating more processing difficulty for the illegal clusters. In the late time-window, illegal clusters generally gave rise to smaller N400s, which would indicate less processing difficulty for the illegal clusters.

Figure 8. Mean differences and SEs between "legal" and "illegal" *ge*-words and *ver*-words in the N400 time window averaged over the three midline channels. Panel A shows the difference between legal and illegal words (averaged across conditions) for each group. Panel B shows the difference between legal and illegal *ge*-words and between legal and illegal *ver*-words averaged across groups. Panel C displays the differences between legal and illegal words split by Group and Condition.



Discussion

The present study investigated how reduction types that differ in their regional distribution are processed by different groups of German listeners. We were interested in the processing of *ge*-reductions, which are very common in Southern Germany but not in Northern Germany, and *ver*-reductions, which are relatively uncommon in all parts of Germany. Canonical *be*-words and nonwords served as control conditions. The nonwords elicited similar N200 and N400 effects in all four groups. These two conditions provided the framework to evaluate how reduced forms are recognized. Are they treated like normal words or like nonwords or in some other way?

There was clear evidence of effects of reduction on the ERPs, which were modulated by the participants' linguistic background. The *ge*-reductions elicited an N200 effect in the two Northern German groups, but not in the Southern German group. The *ver*-reductions elicited more negative difference waves for the three experimental groups in the N200 time window and for the two Northern German groups (but not for the Southern German group) in the N400 time window compared to the control group who heard canonical *ver*-words. That is, as predicted, the groups differed in their processing of *ge*-reductions and, interestingly, there were also group differences in processing *ver*-reductions.

The data also show clear evidence of long-term exposure that eliminates reduction costs. The Southern German group, being very familiar with *ge*-reductions, did not show any processing difficulties for the reduced *ge*-words. While none of the three experimental groups showed an effect in the N400 region for the reduced *ge*-words (i.e., all three groups processed the reduced *ge*-words similarly to the control group, who heard canonical *ge*-words), only the Southern Germans did not show an effect in the N200 region either.

Limited exposure to reductions can help listeners overcome their processing difficulties, though not completely. The two Northern German groups showed N200 effects for both types of reduction, but these processing costs were of short duration, at least for the *ge*-reductions. In the N400 time window, they processed the *ge*-reductions similarly to the control group (who processed canonical forms), while they differed from control when processing the *ver*-reductions. This may be due to limited exposure to *ge*-reductions. In Germany, local coloring is socially accepted[1] so that even by watching the news, Northern

[1] The current minister of Finance, Wolfgang Schäuble, retains a coloring of his Swabian accent in forms of the German word for *is*, which is pronounced /ist/ in Standard German but /iʃt/ in Swabian accents. Schäuble uses predominantly the latter form.

listeners will be exposed to some Southern variants. Apparently, such limited exposure is sufficient to moderate the impact of the reduction costs (as evident in stronger N200 for reduced forms). After this initial exposure, adaptation may reach a plateau, where additional exposure adds little. Even after months and years of living in Southern Germany (M = 2 years 9 months), the migrant Northern Germans still showed an N200 effect for reduced *ge*-words and could not take advantage of their greater experience with this reduction type.

The present study also examined effects of short-term exposure to reductions measured as a change in processing over the course of the experiment. The comparison of first and second experimental halves did not reveal any effects. That is, none of the experimental groups seemed to adapt to reduced *ge*-words or reduced *ver*-words in the short term.

Investigating the effect of phonotactic legality revealed an interesting pattern that was shared by the three experimental groups. While in the N200 time window the reduced *ge*-words were processed the same, regardless of whether they resulted in sequences which are phonotactically illegal in standard German or not, this factor did matter for the reduced *ver*-words. Reduced *ver*-words that resulted in illegal phonotactic sequences evoked an N200 effect compared to the reduced *ver*-words that resulted in legal phonotactic sequences. In the N400 time window, both reduced *ge*-words and reduced *ver*-words showed an N400 effect for the legal sequences compared to the illegal sequences. The effects were stronger overall for the *ver*-words. That is, reductions resulting in illegal phonotactic sequences may cause processing difficulties initially, but these can be overcome quickly so that illegal sequences—contrary to legal sequences—do not cause processing costs later on when the remainder of the reduced word is encountered.

Spinelli and Gros-Balthazard (2007) did not find processing costs for schwa-deletions in the first syllable of French lexical morphemes compared to schwa-intact forms if these deletions resulted in illegal phonotactic sequences, whereas they found processing costs for schwa-deletions resulting in legal phonotactic sequences. Investigating schwa-deletions in inflectional morphemes (*ge*-) and in another language (German), we replicated and extended these findings. The data of the Southern German group fit the findings of Spinelli and Gros-Balthazard. The Southern Germans did not show an N200 effect for *ge*-reductions overall and they processed legal and illegal *ge*-reductions identically in the N200 time window, but they showed processing costs for legal *ge*-reductions compared to illegal *ge*-reductions in the

N400 time window. The findings of the two studies are thus compatible for listeners who are very familiar with a specific reduction type.

However, for participants who were only somewhat familiar with *ge*-reductions, we did not completely replicate the findings of Spinelli and Gros-Balthazard (2007). The *ge*-reductions overall (including those resulting in illegal sequences) elicited N200 effects and thus processing costs for the two Northern German groups. Importantly, there was no difference in processing legal and illegal *ge*-reductions in the N200 time window for these groups. The N200 effect is thus not caused solely by legal sequences. However, we did observe processing costs for *ge*-reductions resulting in legal phonotactic sequences (compared to reductions leading to illegal phonotactic sequences) in the N400 time window for these two groups. With regard to listeners that are not very familiar with a specific reduction type, the ERPs provide us with a more detailed picture and show that reductions, regardless of whether they lead to legal or illegal phonotactic sequences, may cause processing difficulties early on, but that this early difficulty can be overcome fast and probably is then not reflected in behavioral tasks (assuming that the participants in the Spinelli and Gros-Balthazard study were not extremely familiar with schwa-deletions).

The processing costs for reductions leading to legal phonotactic sequences are probably due to increased lexical competition: Legal sequences may activate all words that start with the consonant cluster in question (e.g., the reduced form of *gelacht* [glaxt] may activate all words starting with /gl/) – regardless of whether repair processes such as schwa-epenthesis take place. This would lead to prolonged lexical competition relative to the illegal sequences where similar CC-competitors will not be activated, as they do not exist (or are low in number, e.g., place names).

Another point is that reductions leading to illegal sequences may be overcome faster than reductions leading to legal sequences. First, the uniqueness point of an illegally reduced word (e.g., *gmacht* for *gemacht* 'made') is probably reached earlier than in a legally reduced word. This might already exclude a substantial amount of potential competitors early on. Second, an illegal sequence may trigger a repair process that induces an epenthetic schwa in perception. Dupoux et al. (1999) showed that vowel epenthesis occurs frequently and robustly when Japanese listeners encounter consonant clusters (which are not allowed in Japanese), for example in loan words and nonwords. Vowel epenthesis might not only occur when listeners encounter foreign words containing sequences that are phonotactically illegal in their native language, but also when they hear words from another regional accent of their native language that contain illegal consonant clusters. That is, if a repair process is necessary (i.e., if the reduced form is not stored as such in the mental lexicon), it might be faster for illegal sequences than for legal ones as (at least in our study) not only the sentence context, but also the phonotactics trigger the repair process.

Interestingly, we observed substantial differences in processing *ge*- and *ver*-reductions with respect to phonotactic legality. Contrary to *ge*-reductions, *ver*-reductions do show greater processing costs (as measured by more negative ERPs) when resulting in illegal sequences than when resulting in legal sequences in the N200 time window. Also, the N400 effect for reductions leading to legal sequences was more pronounced for *ver*-words than for *ge*-words. There are at least two reasons that can account for these differences. First, it may be harder for the listeners to recover two deleted segments (in the case of *ver*-reductions) than one (in the case of *ge*-reductions) and this may cause the more pronounced effects for processing the *ver*-reductions. Second, the *ver*-reductions do not occur frequently in natural speech. This might lead listeners up the garden path when hearing legally reduced forms and cause the large N400s across groups. Similarly, when hearing reduced *ver*-words beginning with illegal consonant clusters, which are extremely unlikely to occur in German (e.g., [fs] in reduced *versalzen* 'oversalted' or [ft] in reduced *verteidigt* 'defended'), listeners may show an enlarged N200 response compared to the illegal clusters of the *ge*-reductions, which actually do occur in (Southern) German. The illegal consonant clusters in the reduced *ver*-words are not more complex than the illegal *ge*-clusters in terms of number of consonants the initial cluster consists of (2.7 for *ge*- and 2.5 for *ver*- on average). However, they are more complex in their sequential arrangements. Whereas illegal *ge*-clusters start with a velar (/g, k/) followed by a fricative (/f, ʃ/) or a nasal (/m/), illegal *ver*-clusters start with a fricative (/f/) followed by another fricative (/s, ʃ, v/) or a stop (/b, d, g, p, t, k/). Especially the combination of labio-dental fricative and alveolar fricative or stops is hard to articulate and therefore very uncommon in German.

Comparing the processing of the reduced *ge*-words across groups and thus looking at adaptation in the longer term, we clearly observe a benefit of long-term exposure. Southern German participants who heard reduced *ge*-words did not differ in the way they processed these words from control participants who heard canonical versions of the *ge*-words. This very likely reflects storage of multiple representations. The Southerners may have established schwa-less representations for *ge*-words early on if they got reduced input from their parents, while schwa-bearing representations might have evolved with schooling, exposure to standard German, and learning to read.

The two Northern German groups show striking perceptual flexibility with respect to the processing of the reduced *ge*-words. Both groups show a temporary disruption around 200 ms after target onset. In the N200 time window, they differed in their processing from the control group who heard canonical forms. However, in the N400 time window, they processed the reduced *ge*-words similarly to the Southern and the control group. That is, the initial disruption is easily overcome and the lexicon successfully accessed. Apparently no extra exposure is needed for recognizing reduced *ge*-words, as the non-migrant Northern Germans (having less exposure) did not differ from the migrant Northern Germans (having more exposure to this reduction type).

There are several possible explanations why the two Northern German groups can cope with *ge*-reductions so fast. First, participants in both groups may have stored representations of reduced *ge*-words in their mental lexicon. While this is plausible for participants in the migrant group who had been living in Southern Germany at the time of testing for several months (if not years) and who probably have encountered numerous *ge*-reductions during that time, it seems less likely for participants in the non-migrant Northern German group. These participants had never lived in Southern Germany or Austria and may know *ge*-reductions only from holiday, through the media or from individual Southerners living in Northern Germany. Moreover, it does not seem plausible that participants show a disruption (around 200 ms) for forms that are stored as such in the mental lexicon.

Alternatively, the two Northern German groups may have coped with *ge*-reductions using a very fast prelexical adaptation mechanism. During the first few trials, participants in these groups may have set up or adjusted a mapping rule that assigns the reduced input [g]/[k] to the underlying sequence /gə/. The initial disruption caused by the unexpected consonant cluster may trigger this mapping procedure. A prediction is thus that during the first trials, these participants would show a larger N400 than was measured across all trials, while the N200 would stay the same. Because averaging over only a few trials (e.g., five) will not provide clean ERPs, we cannot test this hypothesis.

Lexical repair is a third explanation for how the two Northern Germans groups could overcome the initial disruption (around 200 ms after target onset) so fast that they did not show any processing difficulties for *ge*-reductions around 400 ms anymore. At a lexical level, these participants would have to insert a schwa into the initial consonant clusters. The initial disruption might, as in the previous account, trigger the repair. In contrast to the prelexical adaptation mechanism, the repair occurs at the lexical level, for each word individually. There is thus no learning and no generalization taking place. According to this view, there should not have been a larger N400 for the first trials.

Comparing the processing of the reduced *ver*-words across groups, we can see that all groups have processing difficulties. None of the three experimental groups processes the (reduced) *ver*-words similar to the control group (who heard canonical *ver*-words) in the N200 time window. However, in the N400 time window, the Southern Germans do not differ from control participants, while the participants from the two Northern groups do. That is, the Southern group shows for the *ver*-reductions a pattern similar to that of the two Northern groups for the *ge*-reductions. Regarding compensation mechanisms, a similar reasoning as for the compensation of *ge*-reductions in the two Northern groups applies. That is, Southern Germans may recover from *ver*-reductions via fast prelexical adaptation or via lexical repair.

Interestingly, the lower processing costs for *ver*-reductions for the Southern listeners compared to the Northern listeners cannot be accounted for by a smaller difference in distance of the reduced form to stored pronunciation variants. As both the Northerners and the Southerners pronounce *ver*-prefixes usually as [fɐ], participants of all three groups probably have stored this variant in their mental lexicon. The more effortless recognition of reduced *ver*-words by the Southern participants can thus more likely be attributed to their greater familiarity in dealing with prefix-reductions in general.

In contrast to the findings for *ge*-reductions, the two Northern German groups differed in their processing of *ver*-reductions. In the N200 time window, the migrant group showed a larger N200 effect than the non-migrant Northern group. This is quite surprising, because if one expected any differences in processing the *ver*-reductions, one would expect these differences in the opposite direction: The non-migrant Northern Germans, who are least familiar with prefix reductions in general, might have more difficulties in processing the uncommon *ver*-reductions and should therefore show a larger N200. However, it could also be that the non-migrant Northern Germans are quite naïve listeners who assume that a Southern German speaker might produce a variety of reductions (i.e., other than *ge*-reductions that they have some familiarity with). So, being less familiar with the Southern German way of speaking, the non-migrant Northerners might accept unusual pronunciations more readily and therefore show a smaller N200 than the migrant Northerners. This hypothesis is supported by the findings of Brouwer, Mitterer and Huettig (2012) indicating that listeners are more tolerant of acoustic mismatches between input and canonical form in a listening situation with casual speech which includes a great deal of reduced forms. In our experiment, participants heard a lot of *ge*-reductions that the non-migrant Northern Germans are apparently somewhat familiar with. Hearing so many reduced forms that actually occur in real world settings might make them more tolerant towards other reduced forms like *ver*-reductions even though they have not had a lot of experience with them. This can also be seen in the processing of the *ver*-reductions that lead to illegal phonotactic sequences. The two Northern German groups also differed in that respect. In the N400 time window, the non-migrant Northern Germans showed a smaller difference between reduced *ver*-words resulting in legal and those resulting illegal sequences than the migrant Northern Germans. While the pattern for *ver*-reductions leading to legal phonotactic sequences is comparable across the three experimental groups, the pattern for *ver*-reductions leading to illegal phonotactic sequences differs in that it is flatter for the non-migrant Northern group than for the migrant Northern and Southern group. This suggests that the non-migrant group may be more tolerant towards these forms.

Viewed from the perspective of the migrant Northern group, the group differences in the processing the *ver*-reductions might suggest another interpretation, and that is that the migrant Northern Germans react in a hyper-sensitive way to *ver*-reductions. They differ in their overall processing of the *ver*-reductions from all other groups, showing the largest N200 effect and therefore the greatest processing difficulties. In the N400 time window, they show a significantly larger difference between reduced *ver*-words resulting in legal and those resulting in illegal sequences than the non-migrant Northern Germans and a numerically larger difference than the Southern Germans (see Figure 8). The migrants may show this relative intolerance to *ver*-reductions because they have just learned that these are *not* part of the Southern German accent. They are not naïve listeners anymore and may have built up representations or mapping rules to deal with reductions occurring in Southern Germany. However, these mechanisms for dealing with reductions may not be so flexible (yet) to suddenly deal with an unknown reduction type. So, in contrast to the Southern Germans, the migrants may not have enough experience with reductions to process the rare *ver*-reductions more easily and, in contrast to the non-migrant Northern Germans, they may have too much experience to be able to simply accept this unusual reduction type.

Regarding regional accents, full adaptation within an experiment has been reported—regardless of whether the regional accent was familiar to the listeners or not (Floccia, et al., 2006). Also, for the supposedly harder case of foreign accents, adaptation has been found to be rapid, even when the accent was unfamiliar (Clarke & Garrett, 2004; Witteman, Weber, & McQueen, 2013). It seems thus surprising that we do not observe an effect of short-term adaptation, especially with regard to the *ge*-reductions. It may be that listeners cannot adapt to a relatively rare reduction type (*ver*-reductions) in the course of a one-hour-experiment (hearing 80 examples of such a reduction) and that to that end a longer consolidation period is needed (cf. effects of the consolidation of newly-learned words; Gaskell & Dumay, 2003). And it seems plausible that the Southern Germans do not need to adapt to *ge*-reductions, as they are very familiar with them. But it is surprising that neither the migrant group nor the non-migrant Northern German group shows an improvement (as measured by lower N200s) in processing the reduced *ge*-words over the course of the experiment. One possible explanation is that the mixed evidence for schwa-reduction (reduction in *ge*-prefixes but not in *be*-prefixes) has had a great influence on the results. However, it is not yet known whether learning about reduced *ge*-participles would transfer to *be*-participles. Moreover, the two prefixes differ in several points such as morphological status (inflectional vs. derivational) and frequency.

Another explanation for not observing a short-term adaptation effect may be that comparing the ERPs of the first experimental half to the second experimental half was too crude a measure. Participants might already show improved processing of *ge*-reductions after 20 items (as suggested by studies on speaker idiosyncrasies as the one by Norris, McQueen, and Cutler (2003) or even after 10 items as found by Kraljic and Samuel (2007)). By comparing experimental halves (i.e., the first 40 occurrences of *ge*-reductions to the second 40 occurrences), we might have failed to observe a learning effect that is actually present. Note, however, that we are constrained here by the ERP method, which requires a large number of trials to generate a reliable ERP (Luck, 2005).

Finally, it could be the case that the two Northern German groups have performed at ceiling when recognizing reduced *ge*-words, which might explain the absence of a short-term adaptation effect. The two groups have limited but apparently sufficient experience with this reduction type and have evolved a compensation mechanism that is good enough to successfully recognize reduced *ge*-words. As *ge*-reductions do not occur in the standard accent nor in their regional accent, and are thus not a native phenomenon for the Northerners, recovering from *ge*-reductions on a within-trial basis might be the best the Northerners can achieve.

Conclusions

The present study revealed processing costs for reductions resulting in legal phonotactic sequences that arise in the N400 time window. These probably reflect increased lexical competition as the reduced forms may activate not only prefix-competitors (through stored representations or schwa-epenthesis, depending on familiarity with the reduction type), but also legal consonant-cluster competitors. Reductions leading to illegal sequences may show processing costs in the N200 time window. These are, however, overcome fast and do not lead to increased lexical competition.

The present study also shows that the degree of familiarity with a specific reduction type influences lexical access. Unsurprisingly, long-term exposure is beneficial for recognizing reduced words and manifests itself in stored representations of reduced variants. Interestingly, listeners with less experience showed great perceptual flexibility and could overcome initial disruptions caused by reduced words quickly—possibly through a lexical repair mechanism. This was also the case if the reduction type was relatively unfamiliar but participants were used to dealing with prefix-reductions. Otherwise, unfamiliar reduction types appear to cause more long-lasting processing difficulties. In this case, moderate experience with reductions in general may even be more challenging for the listener than little experience, as the perceptual system might not yet have adjusted completely and would therefore be unstable.

Overall, the results suggest that familiarity with specific reduction styles plays an important role in processing reduced forms. Speakers of a regional accent develop representations for words reduced according to the norms of their region, while speakers from elsewhere recognize such variants through rapid perceptual adjustments. Furthermore, experience with regionally constrained reductions enhanced listeners' flexibility in recognizing reduced words that do not obey regional or standard pronunciation rules.

Appendix: Critical Stimuli

Table A1. Canonical and reduced form of the critical *ge*-participles and their word frequency per million according to SUBTLEX-DE (Brysbaert, et al., 2011).

Canonical participle	English translation	Frequency	Reduced form	Legal sequence
gearbeitet	worked	3.3	garbeitet	yes
geärgert	annoyed	1.9	gärgert	yes
gehasst	hated	2.5	kasst	yes
geheiratet	married	2.9	keiratet	yes
geheult	cried	1.8	keult	yes
gehofft	hoped	2.9	kofft	yes
geholfen	helped	3.2	kolfen	yes
gehustet	coughed	1.2	kustet	yes
gelächelt	smiled	1.8	glächelt	yes
gelacht	laughed	2.5	glacht	yes
gelandet	landed	2.7	glandet	yes
gelangweilt	bored	2.1	glangweilt	yes
gelästert	blasphemed	1.0	glästert	yes
gelaufen	walked	3.0	glaufen	yes
gelegt	laid	2.8	glegt	yes
gelernt	learnt	3.4	glernt	yes
gelesen	read	3.3	glesen	yes
geliebt	loved	3.1	gliebt	yes
gelobt	praised	2.0	globt	yes
gelogen	lied	2.9	glogen	yes
gelohnt	has been worth it	2.0	glohnt	yes
gelungen	succeeded	2.4	glungen	yes
gerächt	avenged	1.7	grächt	yes
geraucht	smoked	2.4	graucht	yes
gerechnet	calculated	2.3	grechnet	yes
geredet	talked	3.2	gredet	yes
geregnet	rained	1.8	gregnet	yes
gereicht	has been enough	2.0	greicht	yes
gerettet	saved	3.4	grettet	yes
geritten	ridden	2.0	gritten	yes
gerufen	called	2.8	grufen	yes
gerührt	touched	2.2	grührt	yes
gerülpst	burped	0.6	grülpst	yes
gesagt	said	4.4	ksagt	yes
gesehen	seen	4.2	ksehen	yes
gesetzt	set	2.9	ksetzt	yes
gesiezt	formally addressed	0.5	ksiezt	yes
gesucht	searched	3.2	ksucht	yes

gesungen	sung	2.5	ksungen	yes
geübt	practiced	2.2	gübt	yes
gefahren	driven	3.1	kfahren	no
gefangen	caught	3.1	kfangen	no
gefehlt	lacked	2.5	kfehlt	no
gefeiert	celebrated	2.2	kfeiert	no
gefeuert	fired	3.0	kfeuert	no
geflogen	flown	2.7	kflogen	no
gefragt	asked	3.4	kfragt	no
gefreut	has been glad	2.5	kfreut	no
gefroren	frozen	1.9	kfroren	no
gefrühstückt	has had breakfast	1.8	kfrühstückt	no
gefühlt	felt	2.7	kfühlt	no
gefunden	found	3.9	kfunden	no
gejammert	moaned	1.3	gjammert	no
gemacht	made	4.2	gmacht	no
gemalt	painted	2.5	gmalt	no
gemeint	meant	3.1	gmeint	no
gemeldet	reported	2.7	gmeldet	no
gemerkt	noticed	2.8	gmerkt	no
geschehen	happened	3.4	kschehen	no
geschenkt	given as a present	2.7	kschenkt	no
geschickt	sent	3.4	kschickt	no
geschlafen	slept	3.2	kschlafen	no
geschlagen	beaten	3.1	kschlagen	no
geschmeckt	tasted	2.0	kschmeckt	no
geschnarcht	snored	1.4	kschnarcht	no
geschnitten	cut	2.6	kschnitten	no
geschrieben	written	3.3	kschrieben	no
geschwommen	swum	1.6	kschwommen	no
gespannt	tensed	2.4	kspannt	no
gespart	saved	2.2	kspart	no
gespeichert	stored	2.0	kspeichert	no
gespielt	played	3.2	kspielt	no
gesprochen	spoken	3.4	ksprochen	no
gestohlen	stolen	3.3	kstohlen	no
gestorben	died	3.4	kstorben	no
gestört	disturbed	2.6	kstört	no
gestreichelt	caressed	1.4	kstreichelt	no
gestreikt	stricken	1.1	kstreikt	no
gestrichen	cancelled	2.4	kstrichen	no
gestritten	argued	2.3	kstritten	no
Average		2.6		

Table A2. Canonical and reduced form of the critical *ver*-participles and their word frequency per million according to SUBTLEX-DE (Brysbaert, et al., 2011).

Canonical participle	English translation	Frequency	Reduced form	Legal sequence
verlangsamt	decelerated	1.7	flangsamt	yes
verlangt	demanded	3.0	flangt	yes
verlassen	left	3.7	flassen	yes
verlaufen	got lost	2.6	flaufen	yes
verlebt	spent	0.7	flebt	yes
verlegt	lost	2.4	flegt	yes
verleimt	glued together	0.5	fleimt	yes
verleitet	enticed	1.4	fleitet	yes
verlernt	forgotten	1.4	flernt	yes
verlesen	misread	1.6	flesen	yes
verletzt	hurt	3.4	fletzt	yes
verleugnet	disclaimed	1.3	fleugnet	yes
verliebt	fallen in love	3.2	fliebt	yes
verlobt	got engaged	2.4	flobt	yes
verloren	lost	3.8	floren	yes
verlost	raffled	0.5	flost	yes
verrannt	has become obsessed	0.8	frannt	yes
verraten	betrayed	3.2	fraten	yes
verrechnet	miscalculated	1.6	frechnet	yes
verreckt	kicked the bucket	1.6	freckt	yes
verregnet	spoiled by raining	0.3	fregnet	yes
verreist	gone on a journey	2.0	freist	yes
verrenkt	cricked	1.1	frenkt	yes
verrichtet	done	1.4	frichtet	yes
verrieben	put on	0.3	frieben	yes
verrissen	savaged	0.8	frissen	yes
verronnen	trickled off	0.0	fronnen	yes
verrostet	corroded	1.1	frostet	yes
verrottet	rotted	1.5	frottet	yes
verrückt	displaced	3.8	frückt	yes
verrührt	mixed	0.5	frührt	yes
verrußt	sooted	0.0	frußt	yes
verrutscht	shifted	1.2	frutscht	yes
verhaftet	arrested	3.1	fhaftet	no[a]
verheimlicht	kept s.th. secret	2.1	fheimlicht	no[a]
verhext	bewitched	1.8	fhext	no[a]
verhofft	hoped for	0.0	fhofft	no[a]
verholfen	helped	1.6	fholfen	no[a]
verhört	misheard	2.3	fhört	no[a]

verhütet	prevented	0.7	fhütet	no[a]
verbaut	obstructed	1.1	fbaut	no
verboten	forbidden	2.9	fboten	no
verbraucht	spent	2.1	fbraucht	no
verdampft	evaporated	1.3	fdampft	no
verdient	deserved	3.3	fdient	no
verdrängt	pushed aside	1.8	fdrängt	no
verdrückt	skived off	1.6	fdrückt	no
vergessen	forgotten	3.9	fgessen	no
verglichen	compared	2.3	fglichen	no
vergriffen	assaulted	1.3	fgriffen	no
verkauft	sold	3.2	fkauft	no
verkühlt	caught a chill	0.8	fkühlt	no
vermisst	missed	3.2	fmisst	no
verpasst	missed	3.1	fpasst	no
versalzen	oversalted	1.4	fsalzen	no
verschätzt	has been out in his estimation	1.1	fschätzt	no
verschenkt	given away	1.7	fschenkt	no
verschlafen	overslept	1.9	fschlafen	no
verschluckt	choked on	2.2	fschluckt	no
verschont	spared	2.2	fschont	no
verschrieben	miswritten	1.7	fschrieben	no
verschüttet	spilled	2.0	fschüttet	no
verschwunden	disappeared	3.2	fschwunden	no
versetzt	stood s.o. up	2.6	fsetzt	no
verspätet	delayed	2.1	fspätet	no
versprochen	promised	3.3	fsprochen	no
verstanden	understood	3.8	fstanden	no
verstaucht	sprained	1.7	fstaucht	no
verstaut	stowed	1.4	fstaut	no
versteckt	hidden	3.2	fsteckt	no
versucht	tried	3.7	fsucht	no
verteidigt	defended	2.3	fteidigt	no
vertragen	made up	2.5	ftragen	no
vertraut	trusted	3.0	ftraut	no
vertröstet	put s.o. off	0.6	ftröstet	no
verwählt	misdialed	1.6	fwählt	no
verwirrt	confused	2.9	fwirrt	no
verwöhnt	regaled	2.0	fwöhnt	no
verzählt	miscounted	1.1	fzählt	no
verziehen	forgiven	2.2	fziehen	no
Average		1.9		

[a] Treated as a legal sequence in the analyses.

Table A3. Canonical form of the critical *be*-participles and their word frequency per million according to SUBTLEX-DE (Brysbaert, et al., 2011) and the nonwords.

Canonical participle	English translation	Frequency	Nonword	Legal sequence
behagt	liked	1.1	pugt	yes
behalten	has been proven correct	3.4	pilten	yes
behandelt	treated	3.0	pindelt	yes
behängt	draped	1.0	pöngt	yes
beharrt	insisted	0.9	pirrt	yes
behauen	worked (stones)	0.0	piehen	yes
behauptet	claimed	2.8	peiptet	yes
behaust	housed	0.0	peist	yes
beheizt	heated	1.0	pauzt	yes
behelligt	molested	0.6	polligt	yes
beherbergt	accommodated	1.3	porbargt	yes
beherrscht	dominated	2.3	parscht	yes
beherzigt	heeded	1.0	parzigt	yes
beherzt	courageous	0.6	pirzt	yes
behext	bewitched	0.7	paxt	yes
behindert	hampered	2.3	pondert	yes
behoben	repaired	1.6	piben	yes
beholfen	managed with s.th.	0.0	paulfen	yes
behütet	protected	1.4	pietet	yes
belächelt	smiled at s.th.	0.0	blüchilt	yes
beladen	loaded	2.0	blidden	yes
belagert	besieged	1.4	bliegert	yes
belangt	prosecuted	1.0	blengt	yes
belassen	left it at s.th.	2.1	blüssen	yes
belastet	strained	2.1	blestet	yes
belästigt	bothered	2.4	blostegt	yes
belauert	stalked	0.8	bloërt	yes
belaufen	added up to	0.9	bliffen	yes
belebt	stimulated	1.4	bluhbt	yes
belegt	occupied	2.2	blugt	yes
belehrt	taught	1.3	blohrt	yes
beleidigt	affronted	2.6	blodigt	yes
beleuchtet	illuminated	1.8	blachtet	yes
belichtet	exposed to light	0.7	blochtet	yes
beliebt	popular	2.6	bläbt	yes
beliefert	supplied	1.2	blaufert	yes
beliehen	lent on s.th.	0.0	blahen	yes
belobigt	commended	0.5	blabigt	yes
belogen	told a lie	2.4	bleugen	yes

belohnt	rewarded	2.1	blient	yes
belüftet	aired	0.3	bläftet	yes
belustigt	amused	0.6	blistugt	yes
beraten	advised	2.1	breten	yes
beratschlagt	deliberated	0.3	brittschlogt	yes
beraubt	bereaved	2.1	breibt	yes
berauscht	inebriated	1.3	brischt	yes
berechnet	calculated	1.9	brachnit	yes
berechtigt	justified	1.9	bruchtegt	yes
bereichert	enriched	1.3	brüchert	yes
bereinigt	sorted out	1.5	branegt	yes
bereist	visited	1.3	bräst	yes
bereitet	caused	2.7	brietet	yes
bereut	regretted	2.0	brött	yes
berichtet	told	2.4	brochtet	yes
berichtigt	corrected	1.1	brochtigt	yes
berieselt	sprinkled	0.3	bruselt	yes
berücksichtigt	accounted for	1.7	bracksochtigt	yes
berufen	appointed	2.0	brofen	yes
beruhigt	calm down	2.8	brehigt	yes
berührt	touched	2.7	brährt	yes
beruht	been based on	2.1	breht	yes
besagt	proved	2.3	pseigt	yes
besänftigt	appeased	1.3	psonftigt	yes
besäuselt	tipsy	0.3	psüselt	yes
besehen	looked at	0.6	psohen	yes
beseitigt	removed	2.2	psautigt	yes
besessen	owned	2.7	psussen	yes
besetzt	occupied	2.5	psutzt	yes
besichtigt	visited	1.1	psachtigt	yes
besiedelt	populated	0.8	psäudelt	yes
besiegelt	sealed	1.8	psaugelt	yes
besiegt	defeated	2.7	psugt	yes
besoffen	got plastered	2.2	psiffen	yes
besohlt	resoled	0.0	psühlt	yes
besoldet	salaried	0.3	pseldet	yes
besonnen	bethought	1.4	psennen	yes
besorgt	procured	3.1	psergt	yes
besucht	visited	2.8	psecht	yes
besudelt	besmirched	1.3	psedelt	yes
besungen	sung about	0.5	psöngen	yes
Average		1.5		

Summary and conclusions

Chapter 6

This thesis combined two previously independent lines of research that both address how listeners overcome adverse conditions in spoken word recognition: research on phonological reductions and on talker adaptation. The reason for combining the two was that, because speakers differ in their tendency to produce phonological reductions, it would be useful for listeners to be able to adapt to the "reduction style" of individual speakers. Looking at whether and how listeners adapt to speaker-specific reductions also has implications for the debate on the structure of the mental lexicon. It may be that listeners predominantly store instances of reduced words that they encounter in an episodic lexicon. One should then find strong evidence for word-specific learning. However, if listeners mostly use abstraction processes to recognize reduced words in a phonologically abstract lexicon, one should observe generalization of learning to new words that are reduced in the same way. Furthermore, if listeners can abstract over the reduction patterns they encounter, the question arises how much degradation of the speech signal they can handle. This thesis thus asked whether listeners are able to deal only with segmental reductions or also with reduction of larger linguistic units, such as the syllable. Another issue this thesis addressed was the time course of adaptation to reductions. I looked at two aspects of the time course: First, at what point in time is learning applied when recognizing a reduced word and, second, how much experience is necessary to adapt to a specific reduction pattern.

Summary of the results

Examination of the amount of exposure necessary for learning to occur and of the underlying learning function may reveal possible differences between different forms of adaptation. Learning may arise incrementally and reach a plateau at some point or it may occur in a step-wise fashion (i.e., "now I get it"). However, even for the well established case of learning about speaker idiosyncrasies (e.g., McQueen, Cutler, & Norris, 2006; Norris, McQueen, & Cutler, 2003), there was no previous work that clearly indicated when and with what time course learning occurred. Therefore, one of the first steps in this project was to investigate the time course of perceptual learning about ambiguous fricatives using the eye-tracking paradigm. This learning process is potentially very similar to adaptation about segmental reductions. It was thus hoped that examining the time course of learning could provide useful information for further investigations about reduction.

In the experiment reported in Chapter 2, the goal was to determine how many tokens are necessary for learning to occur and what the learning function looks like. Test trials were therefore interspersed among training and filler trials. In the training trials, two groups of Dutch participants were exposed to [$^{s}/_{f}$]-bearing words and learned to interpret this ambiguous sound as either /s/ (in the /s/-bias group) or /f/ (in the /f/-bias group) based on lexical bias. That is, they heard this odd sound at the end of either /s/-final or /f/-final words and had to click on the orthographic form of that word (e.g., on *radijs* ‘radish’ or *olijf* ‘olive’). In the test trials, participants heard the ambiguous sound at the end of a word that could end in either /s/ or /f/ (e.g., spoken *doo[$^{s}/_{f}$]* could refer to either written *doos* ‘box’ or written *doof* ‘deaf’). Learning arose after exposure to 10 training items and 10 temporary minimal pairs (i.e., items like *gister* ‘yesterday’ and *giftig* ‘toxic’ that were still ambiguous when listeners heard the ambiguous sound replacing the fricative). I thus observed similar results to those in Kraljic and Samuel (2007), where learning was also found after exposure to 10 training items. Once the learning effect had emerged, it stayed stable throughout the experiment and did not get stronger with additional training. Learning thus seems to occur in a step-wise fashion. Interestingly, the learning effect arose only late within a trial (ca. 900 ms after target onset). Mitterer and Reinisch (in press) investigated this issue further using eye-tracking in a blocked exposure-test design. They found that perceptual learning takes effect as soon as it possibly can, that is 200 ms after the onset of the fricative. It is thus likely that the interspersed test trials (especially the early ones) interfered with learning.

As the design with test trials interspersed among training trials turned out to be potentially problematic for observing learning effects, a classical design, in which exposure and test trials were blocked, was used in Chapters 3 and 4. The issue of how many items are necessary for learning to occur was not pursued further. Instead, I focused on perceptual learning of reductions.

The aim of Chapter 3 was to test whether listeners can adapt to segmental and syllabic reductions in continuous speech. Critically, this chapter focused on generalization of learning. Two groups of Dutch participants were exposed to different reduction types and it was tested whether this helped them to recognize other words that were reduced in a similar way. Such a generalization would show that listeners make use of prelexical abstraction.

The two reduction types under investigation were the weakening of a word-initial /b/ to a labio-dental approximant ([ʋ]) and the shortening of the prefix *ver-* to a long [fː]. These two reduction types differed in the size of the reduced unit (segment vs. syllable), so that adaptation could inform us about the size(s) of abstract units represented at the prelexical level. If learning occurs for both types of reductions, it would imply that prelexical abstraction can make use of units that are larger than a single segment. The design was such that a segmental reduction group was exposed to reduced /b/-words, while they encountered *ver*-initial words in canonical form. A syllabic reduction group was exposed to reduced *ver*-words, while they encountered /b/-initial words in canonical form. In this way, the two groups served as control groups for each other. In the test phase, both groups heard both types of reductions in new words and I compared how well they recognized these reduced words.

In Experiment 3.1, a learning effect was observed for the syllabic reduction group, but not for the segmental reduction group. That is, the syllabic reduction group had a greater target preference for the *ver*-reductions than the segmental reduction group, but the segmental reduction group did not have a greater target preference for the /b/-reductions than the syllabic reduction group. This null effect could have been due to one of the many aspects concerning the creation of the stimuli. The one that seemed most likely to me was an inappropriate reduction context: The speaker devoiced the schwa preceding the reduced targets with the consequence that the /b/ did not occur in intervocalic position. This intervocalic context, however, might be crucial for /b/-reductions to occur. For that reason, the carrier sentence was changed in a follow-up experiment. In Experiment 3.2, learning effects were then observed for both the syllabic reduction group and the segmental reduction group, showing that the phonetic context is important in determining how listeners process reduced forms.

As in Chapter 2, the learning effects are reflected late in the eye-movement data. Participants showed a greater preference for the reduced target only from about 600-800 ms after target onset.

Since the reduced test items had not been encountered in the exposure phase, learning about /b/- and *ver*-reductions generalized across words. If learning takes place at a lexical or at a postlexical level, generalization of learning should not be observed. If, however, learning takes place at a prelexical level, it should automatically apply to all words showing the same reduction pattern. The fact that learning about sublexical units was applied to other words suggests that mechanisms compensating for segmental and syllabic reductions take place at a prelexical level. The results suggest thus that an abstractionist mode of processing is involved in lexical access. However, current abstractionist models need to be revised, as they do not include representations larger than a segment at the prelexical level (which are required to account for adaptation to *ver*-reductions). It has to be noted, however, that the syllabic reduction investigated in Chapter 3 was also the reduction of a morpheme. It hence remains unclear whether the larger-than-segment representations at the prelexical level are restricted to morphemes.

Furthermore, the learning effects found in Chapter 3 were not strong. Two scenarios can account for this. First, it may be that listeners can apply their knowledge about reductions not only to the reduction type they are familiar with, but also to other reduction types. In this case, both groups of listeners handle /b/- and *ver*-reductions very well with a small advantage for the more familiar reduction type. The second, opposing scenario would be that learning about reductions is predominantly word-specific. Both groups of listeners handle /b/- and *ver*-reductions then quite poorly, as there is only minimal generalization within and across reduction types.

Chapter 4 addressed these issues. To examine the possibility of generalization across reduction types, a control group, who heard no reductions at all, was introduced next to two experimental groups, who were, as in Chapter 3, exposed to segmental or syllabic reductions respectively. If generalization does occur, I would expect that both experimental groups outperform the control group on both types of reductions. To address the issue of word-specificity, words were repeated in the exposure and test phases. If learning turns out to be mainly word-specific, learning effects should be much larger when words are repeated than when they are not.

In Experiments 4.1 and 4.2, the reduced words of the test phase had already occurred in the exposure phase. The segmental reduction group was exposed to /b/-initial words, in which the onset consonant was reduced to a bilabial nasal, while the syllabic reduction group was exposed to CVC-initial words, in which the unstressed full vowel was deleted. The full-vowel deletions were chosen to investigate further whether the morphemic status of the syllables involved in reduction processes influences the adaptation process. The syllabic reductions examined in Chapter 3 were reductions of a prefix and therefore their canonical forms showed an extremely high consistency in their segmental structure (i.e., it was always the same phoneme that occurred in a certain position). Chapter 4 tested whether listeners can also adapt to the reduction of non-morphemic syllables that necessarily show greater variation in their segmental structure than prefixes. The performance of the two experimental groups was compared to the performance of the control group and not to each other. Learning effects for repeated words were observed for both the segmental and the syllabic reduction group. To test whether these effects were truly word-specific and not due to abstraction processes, Experiment 4.3 was run, which, like Chapter 3, focused on generalization of learning. Learning about vowel deletions did not generalize to other vowel-deleted words. The corresponding learning effects in Experiments 4.1 and 4.2 were thus indeed word-specific. Learning about /b/-reductions, however, generalized to new reduced /b/-words. The corresponding learning effect found in Experiment 4.1 was thus not word-specific. An important difference between the two reduction types that might explain the different kinds of learning processes observed lay in the consistency of the input. While the vowel deletions showed a very varied reduction pattern, the pattern was highly consistent for the /b/-reductions. It thus seems that a consistent reduction pattern facilitates abstraction processes, while varied reduction patterns are dealt with by storing instances of individual reduced words.

Generalization of learning across reduction types was found only for the syllabic reduction group. Participants in this group could use their knowledge about vowel deletions to recognize reduced /b/-words better. For the generalization of learning about /b/-reductions to vowel deletions, however, inhibitory effects were observed (i.e., the segmental reduction group had a smaller target preference for the vowel-deleted words than the control group). The most likely explanation for this asymmetric finding lies again in the inconsistency of the reduction pattern. Listeners who are exposed to a variable reduction pattern seem to become more tolerant towards any form of reduction.

Facilitative effects for word-specific learning and for learning that generalized within reduction type occurred late within a trial (at around 1100 ms after target onset). Inhibitory effects and facilitative effects for generalization of learning across reduction types, however, appeared early (from 200 ms onwards, as in Experiment 4.2).

The different kinds of learning effects observed suggest that listeners use various means to recognize reduced words: They store encountered reductions (the word-specific learning effect) and build up abstract mapping rules in order to apply their knowledge to new reduced words (generalized learning within reduction type) if that is possible. Generalized learning seems to be constrained by the consistency of the reduced input. In the case of non-morphemic syllable reductions, the input was too variable to abstract over it. Instead, the perceptual system then displayed greater flexibility in tolerating other forms of reduction (generalized learning across reduction types).

Chapter 5 returned to the issue of how much experience is necessary for adaptation to take place, which was also addressed in Chapter 2, but was now investigated on a different time scale. Rather than asking how many instances a listener needs to encounter to be able to adapt, this chapter addressed differences in processing due to a varying degree of experience with a given pattern. This was achieved with a quasi-experimental approach in which experience was not manipulated experimentally, but groups of participants with different amounts of experience (little experience with *ver*-reductions; little vs. moderate vs. great experience with *ge*-reductions) were recruited. Furthermore, short-term adaptation was again investigated. It was examined whether listeners, especially those in the less experienced groups, get better in recognizing reduced words over the course of the experiment.

By means of ERPs, long-term and short-term adaptation to two types of German prefix-reductions was investigated. The deviation of reduced words from canonical pronunciation may provoke phonological mismatch detection and therefore elicit an N200 effect. It may also cause difficulties in accessing the mental lexicon and later in integrating the word into the sentence context and therefore elicit an N400 effect. The greater these effects, the more difficulties listeners have in recognizing the reduced forms. One of the two reduction types under investigation occurs very frequently in Southern Germany and Austria (reduction of *ge*- to [g]/[k]), whereas the other is relatively rare in all parts of Germany and Austria (reduction of *ver*- to [f]). A Northern German group, a migrant group (made up of participants who had moved from Northern to Southern Germany) and a Southern German group heard sentences containing these reductions, while a control group heard these sentences with only unreduced forms. Both reductions led to illegal phonotactic sequences in half of the cases. Reductions resulting in illegal phonotactic sequences may cause early processing difficulties, but these were overcome fast—unlike the continuing processing costs caused by reductions leading to legal sequences. The higher processing costs for "legal" reductions are probably due to increased lexical competition, as the reductions leading to legal sequences may activate not only competitors starting with the same prefix, but also competitors starting with the same legal consonant cluster.

For *ge*-reductions, the results showed a benefit of long-term exposure for the Southern listeners and fast perceptual adjustments for the Northern and migrant listeners who were at least somewhat familiar with *ge*-reductions. For *ver*-reductions, the Southerners showed perceptual flexibility, whereas the Northerners and especially the migrants had persistent processing difficulties. The ease with which the Southerners processed *ge*-reductions suggests that they have stored these reduced forms in their mental lexicon. The Southerners also showed less processing difficulties for *ver*-reductions compared to the Northern and the migrant group. As all three groups were relatively unfamiliar with *ver*-reductions, it seems likely that the Southerners took advantage of their greater experience with prefix-reductions in general. While the Northern and the migrant groups were quite unfamiliar with *ver*-reductions, their limited experience with *ge*-reductions was apparently enough for them to be able to flexibly adjust their perceptual system. They showed difficulties with the *ge*-reductions in the N200 time range, but could overcome these fast. Hence, in the N400 time range, they processed these reductions in a "native-like" manner (i.e., like the Southern listeners). Familiarity with a given reduction type is thus necessary for effortless recognition of reduced words.

In contrast to the eye-tracking experiments in Chapters 3 and 4, no short-term adaptation (i.e., adaptation over the course of the experiment) was observed.

Perceptual adjustments, however, took place early on, as reductions that caused processing difficulties in the N200 time window did not do so anymore in the N400 time window.

Conclusions

The experimental chapters of this thesis shed light on specific aspects of the adaptation process to segmental and syllabic reductions. There was evidence that different kinds of mechanisms are involved when listeners adjust to reduced speech. This has consequences for models of spoken word recognition. The results further suggest that adapting to reductions is not an entirely automatic process but rather is constrained by certain factors. Regarding its time course, adaptation to reductions seems to differ from perceptual learning about speaker idiosyncrasies. The following sections will discuss these issues in greater detail.

Processes involved in adaptation to reductions

Different learning mechanisms appear to be involved in the recognition of reduced word forms. Experiments 4.1 and 4.2 (on word-specific learning) and Experiment 5 (which showed no processing difficulties for very familiar forms of reduction) suggested that listeners store reduced pronunciation variants. Experiments 3.1, 3.2 and 4.3 indicated that abstraction processes play an important role in recognizing previously unheard reduced forms of the same reduction type. Finally, the handling of words that were reduced in an unexpected way (new reduction types in the test phases of Experiments 4.1 and 4.3 and the only somewhat familiar *ge*-reductions for the Northern German participants in Chapter 5) provided evidence for the flexibility of the perceptual system.

These different mechanisms possibly complement each other over different situations. Storage of reduced forms is beneficial for later encounters of the same word reduced in the same way, especially if a given reduced form is encountered frequently. In contrast, generalization of learning about reductions helps listeners to recognize new reduced words that were not encountered before. Generalization processes, however, seem to be limited by factors such as input consistency (for generalization within reduction types at the very least). The results of Chapter 4, and especially the inhibitory effects, suggest that generalization from a consistent reduction pattern to a variable reduction pattern is not possible. It might still be the case, however, that learning can transfer from a consistent pattern to another consistent reduction pattern. This seems plausible for very similar reduction types. Future research would have to test whether a strong generalization effect across reduction types can be found for reductions that are similar in size and type (e.g., schwa-deletion in the German prefixes *be-* and *ge-* or reduction processes that affect different stop consonants in a similar way). Whether learning may then also generalize from one consistent reduction pattern to another consistent reduction pattern that differs considerably (e.g., the /b/-reductions and the *ver*-reductions examined in Chapter 3) still remains an open question.

When new reduced forms are encountered that are reduced in an unexpected way, the perceptual system might be flexible enough to adjust to them—if it has had sufficient input beforehand. This prior input likely has to be variable enough so that the system tolerates any form of reduction (e.g., like the vowel deletions in Chapter 4; participants exposed to those could then handle /b/-reductions quite well) and/or similar enough to the unexpected reduced forms. The relatively uncommon reduction of the prefix *ver-* (mostly pronounced as [fɐ]) to [f] investigated in Chapter 5 was similar to the reduction of the prefix *ge-* to [g]/[k] in that a vowel was deleted in both cases. However, only the Southern listeners, who had no difficulties in recognizing the *ge*-reductions, could also handle these *ver*-reductions relatively well. Having sufficient experience with *ge*-reductions (and very likely also other forms of prefix-reductions) and the relative similarity of the two reduction types likely facilitated this perceptual adjustment.

It might seem that only *one* way of adapting to reductions is more efficient than having several different ways to adapt. However, there is apparently not *one* mechanism that can deal with all the different kinds of reductions (mild, severe, consistent, inconsistent, ...). The most economical way to deal with inconsistently and/or strongly reduced word forms may be to just store these "irregular" forms. In contrast, it may be most parsimonious to use abstract mapping rules when dealing with consistently reduced forms. It still has to be investigated how independent these different mechanisms are. It is not clear, for example, whether there can be perceptual adjustments or abstraction mechanisms without there also being stored exemplars over which generalizations can be made (in a more or less specific way).

Constraints on the adaptation process

The above mentioned adaptation processes are constrained by several factors. First, as already discussed, consistency of the input plays a crucial role in determining how listeners adapt to reduced forms. Second, the failure to find generalized learning for /b/-reductions in Experiment 3.1 indicates that the right articulatory context (e.g., intervocalic position) needs to be present for adaptation to certain reductions to occur. Listeners are sensitive to such probabilistic constraints (Mitterer & McQueen, 2009b) and do not adapt if these are not met.

Furthermore, Chapter 5 demonstrated that familiarity with a given reduction type is necessary for effortless recognition of reduced forms. In Chapters 3 and 4, it was verified that the selected reduction types actually occurred in natural speech. Participants had thus likely encountered them before and were at least somewhat familiar with them. The failure to find adaptation to unfamiliar forms of reduction (the *ver*-reductions in Chapter 5 for the Northern and migrant group) suggests that a minimum of exposure is necessary before adaptation can occur.

Moreover, it seems puzzling that some of the learning effects observed in Experiment 4.1 could not be replicated in the later experiments of Chapter 4. While there was at least a trend for generalization of learning from vowel deletions to /b/-reductions visible in both the accuracy data and the eye-tracking data in Experiment 4.3, no such hints were found for an effect of repeated /b/-reductions in Experiment 4.2. Intuitively, listeners should find it easier to recognize a reduced word when they encounter it a second time, regardless of whether they use storage or abstraction processes for recognition. The null effect for repetition of reduced /b/-words in Experiment 4.2 suggests thus that adaptation to reductions does not take place automatically—contrary to what has been claimed for perceptual learning about speaker idiosyncrasies (McQueen, et al., 2006) and adaptation to foreign-accented speech (Witteman, Bardhan, Weber, & McQueen, in press). An aspect that speaker idiosyncrasies (here: the odd pronunciation of a fricative) and foreign-accented speech (here: mispronounced Dutch vowels) share with the syllabic reductions—and that is potentially not present to the same degree in weakened segments—is saliency. Listeners are very likely aware of segments that are missing at the beginning of a word (as they do in the syllabic reductions under investigation) and of segments that are not pronounced in a native-like way. An ambiguous fricative that may represent a speaker's lisp is also very noticeable. In contrast, the deviation of a weakened segment from the canonical pronunciation might not be salient or troublesome enough to warrant adaptation—unless other factors boost its saliency.

One such factor may be predictability. Being able to predict the upcoming and potentially reduced word likely helps listeners to adapt to reductions. In a similar vein, after failing to observe a perceptual learning effect for ambiguous fricatives in word-initial position for words that were presented in isolation, Jesse and McQueen (2011) proposed that lexical knowledge has to be available as the ambiguous sound is being processed for learning to occur. When lexical knowledge was provided during exposure by displaying the orthographic form of a reduced word on the screen, a (generalized) learning effect was observed (see Experiments 3.1, 3.2, 4.3). The printed words on the screen likely help native listeners to adapt to reductions in a way that is similar to that in which foreign subtitles help non-native listeners when watching a movie in that foreign language (Mitterer & McQueen, 2009a). In Experiment 4.2, however, the semantically rich exposure sentences alone were not enough to trigger a learning effect for the repeated /b/-reductions. One reason for this might be the high number of low-frequent words and the difficulty of creating highly predictable sentences for those words. In real-world communication, the wider discourse context helps to increase the predictability of upcoming words and might boost participants' ability to adapt to reduced words. A future experiment could test this hypothesis by examining whether a clear learning effect for repeated high-frequency words that are reduced in a consistent way can be found if they occur in highly predictive sentences or even in wider discourse contexts in the exposure phase. It thus seems plausible that the activation of possible lexical candidates before the reduced signal is being processed is beneficial for the learning process. Making segmental reductions more salient by increasing the predictability of their canonical form likely provokes prediction errors which in turn might lead to implicit learning (Clark, 2013; Pickering & Garrod, 2013).

In conclusion, there is no simple story to be told about when adaptation to reductions may succeed. Whether listeners can adapt to reduced word forms depends on the interplay of several factors, such as input consistency, appropriate articulatory context, familiarity, saliency and predictability. Another possible factor that might influence the adaptation process and that was not studied in this thesis is the consistency with which the speaker produces reductions. In the exposure phases of the experiments in this thesis, a speaker was modeled who reduced all words complying with the conditions for the occurrence of a specific reduction (i.e., words that contained unstressed /b/s but were unreduced were excluded from the experimental materials of Chapters 3 and 4). The question then arises whether listeners can also adapt to reductions if they hear a speaker produce, for instance, /b/-reductions only occasionally. Witteman, Weber, and McQueen (in press) investigated a similar matter by testing whether native listeners can adapt to a (non-native) speaker whose pronunciations were only sometimes accented and passed as native-like otherwise. They found that listeners did adapt, but that they took longer to do so than listeners who were exposed to only accented words.

The time course of adaptation to reductions

Apart from investigating under which circumstances what kind of adaptation mechanisms can be observed, this thesis also focused on when learning becomes evident. Chapter 5 did not reveal any effects of short-term adaptation, but provided time-course information on the processing of reduced forms. Unless reduction types are extremely frequent (which they were not in Chapters 3 and 4), reduced words elicit an N200 effect. An overall N400 effect was only observed for very uncommon forms of reductions in listeners unfamiliar in dealing with the reduced forms. However, reductions leading to legal phonotactic sequences elicited a more negative N400 than reductions leading to illegal phonotactic sequences. The former cause thus more longer-lasting processing difficulties and exactly these legally reduced forms were investigated in Chapters 3 and 4 using eye-tracking paradigms.

The results of these eye-tracking experiments show that the point in time when learning about reductions takes effect may depend on the specificity of the learning process. Facilitative and inhibitory generalization effects across reduction types, which—in Chapter 4—are likely not to be specific to any segments or words, were observed early (from 200ms after target onset onwards). In contrast, the word-specific effect found in Experiment 4.1 was quite late (starting at 1100 ms after target onset). The effects for generalization within reduction type observed in Chapters 3 and 4 occurred similarly late. However, competitor preferences, if they occurred, were observed much earlier, sometimes as early as 300 ms after target onset. This seems to suggest that the bottom-up information in the speech signal has a strong influence on speech recognition (e.g., Marslen-Wilson, 1987) and learning needs time to override the signal-driven interpretation (i.e., “what-you-hear-is-what-you-get”).

This seems plausible for word-specific learning, especially if the reduced word closely resembles another canonical word. At some point in time, the activation of *Parijs* 'Paris' has to win over the activation of *prijs* 'price' when hearing the reduced speech input *P'rijs*. Initially, the activation of *prijs* is likely to be stronger as this meaning is encountered much more frequently. Speaker-specific information (e.g., on the tendency of this speaker to reduce words like *Parijs*) has then to kick in and shift perception in favor of the candidate *Parijs*. This may take some time (i.e., may not happen immediately).

The effects of rule-based learning, however, do not appear to be necessarily delayed. Mitterer and Reinisch (in press), who conducted a follow-up study to Chapter 2 of this thesis, found early learning effects in their eye-tracking data. They concluded that perceptual learning is indeed perceptual rather than post-perceptual once it is complete. The late effects of Chapter 2 can thus be attributed to the interference of the interspersed test trials. But what of the late effects in Chapters 3 and 4?

Mitterer and Reinisch (in press) observed early learning effects using an eye-tracking paradigm with a classical exposure-test design. It thus seems unlikely that the late effects found in Chapters 3 and 4 are caused by the combination of a printed-word eye-tracking task and a learning paradigm. Furthermore, the participants in the study by Mitterer and Reinisch (in press) had to make conscious metalinguistic judgments about the speech sounds they heard by clicking on printed words whose spoken realizations deviated from the canonical pronunciations. As this task did not affect the timing of the effects in Mitterer and Reinisch (in press), it seems unlikely that it should have done so in the experiments presented in Chapters 3 and 4.

If perceptual learning should and does in general occur early, and methodological issues have been ruled out as possible causes, what might then be the reason for the late learning effects observed in this thesis? One point on which Chapters 3 and 4 differ from the study by Mitterer and Reinisch (in press), and other classical perceptual learning studies (e.g., McQueen, et al., 2006), regards the mapping rules listeners must generate or adjust to correctly recognize the deviant sound(s). In the classical perceptual learning paradigm, listeners have to learn to include a previously unknown ambiguous sound into *one* of two possible phoneme categories. They must thus generate a new mapping rule for that sound, and that sound is supposed to be truly ambiguous. Ideally, it is judged as belonging to one of the two categories in 50% of the cases. Apart from the pretested [$^{b}/_{\upsilon}$] sound in Experiment 3.1, none of the reduced sounds used in Chapters 3 and 4 were truly ambiguous. The naturally occurring [$^{b}/_{\upsilon}$]s used in Experiment 3.2, as the replacements for /b/, were likely more /ʋ/-like than /b/-like. The long [fː]s replacing *ver*-s in Experiments 3.1 and 3.2 were clearly more /f/-like than *ver*-like. The equivalent holds for the natural instances of [m] that replaced /b/s in Chapter 4 and also for the CVC-words in that chapter, because the whole vowel was deleted, leaving only the consonant clusters.

In Chapters 3 and 4, listeners thus learned that a particular speaker was likely to pronounce, for example, a /b/ as an [m] and hence that an existing sound (e.g., [m]) mapped onto *two* categories for that speaker (i.e., /m/ and /b/). Their perception of an [m] might therefore have shifted from judging it as /m/ in most cases to judging it as /m/ in 80% and as /b/ in 20% of the cases. With this kind of learning, the initial signal-driven hypotheses still strongly favor the canonical form, and only when later-arriving segments rule that form out can the learning take effect. Therefore, as soon as listeners receive evidence that a particular sound can map onto more than one category, the learning process likely needs more time to take effect.

Studies on foreign-accented speech have also addressed the issue of many-to-one mappings. Hanulíková and Weber (2012) investigated the effect of linguistic experience on perception of different substitutes (/t/, /s/, and /f/) for English /θ/ by Dutch and German listeners. While German listeners preferred the /s/ variants, Dutch listeners preferred the /t/ variants, which is consistent with their experience and their own productions. Although the /f/ variants were perceptually most similar to realizations of /θ/, neither Dutch nor German listeners preferred them.

Some studies on many-to-one mappings in foreign-accented speech also addressed adaptation effects in native listeners. However, these concentrated mostly on the adaptation process itself and not on its time course. Eisner, Melinger, and Weber (2013), for instance, examined perceptual learning about devoiced stop consonants in native English listeners. The participants heard a Dutch speaker produce final devoicing in her pronunciations of word-final /d/s (i.e., she pronounced

/d/s as [t]s). After exposure to these variants, British English listeners were primed by similarly mispronounced words in a cross-modal priming task. That is, they adapted to these variants and recognized the deviant (devoiced) stop as the intended voiced stop /d/. Although cross-modal priming measures word recognition online, it does not give insight into the fine temporal structure of the recognition process. It can only show that adaptation has occurred, but cannot provide information on when exactly it takes effect, that is, whether it occurs as soon as the deviant sound is being processed or only some hundred milliseconds thereafter.

Witteman, Bardhan, et al. (in press) investigated adaptation to Hebrew-accented Dutch also using the cross-modal priming paradigm. The Hebrew speaker pronounced Dutch words containing the short vowel [ɪ] correctly, but shortened the Dutch vowel [i] to [ɪ]. Facilitatory priming was observed in Dutch listeners for the variant words after a short exposure of 3.5 minutes to this accent, in which participants were asked to monitor for a phoneme that was not affected by the accent. It was concluded that adaptation to foreign-accented speech is rapid and automatic.

While research on foreign-accented speech confirms that listeners can adapt to sounds that map onto more than one category, this thesis presents the first studies that also look at the time course of such an adaptation process. Unlike adaptation to reduced speech, adaptation to foreign-accented speech seems to be automatic, just as perceptual learning about speaker idiosyncrasies. However, future research on the time course of adaptation to foreign-accented speech may reveal that, because of the many-to-one mappings involved, learning needs time to take effect. Adaptation to foreign-accented speech might in this respect be more similar to adaptation to reduced speech than to perceptual learning about speaker idiosyncrasies.

Models of spoken word recognition

What do the different forms of adaptation to reductions observed in this thesis tell us about how words are stored in the mental lexicon and, consequently, how they are recognized? Throughout this thesis, there was evidence both for storage of reduced forms in the mental lexicon (Chapters 4 and 5) and for abstraction over the reduced speech signal facilitating generalization effects (Chapters 3 and 4). Reduced words seem to be stored as variant pronunciations in at least two cases. First, they are stored if they show such a varied reduction pattern that no abstraction rule can be generated (e.g., the variable vowel deletions in Chapter 4). But also reduced words that follow strictly regular patterns may be stored as variants in the mental lexicon, if listeners are very familiar with them (e.g., the *ge*-reductions in Chapter 5 which are very familiar to Southern German listeners). Possibly, this claim has to be constrained further so that regular reduced forms are only stored if a) they are also the listeners' native pronunciation of those words (i.e., the reduced forms are acquired before canonical forms are) or b) listeners produce them themselves naturally (i.e., the reduced forms are treated as native pronunciations but are acquired after the canonical forms). Future research will have to test the potential effects that age of acquisition and production have on whether regularly reduced word forms are stored. Furthermore, it seems plausible that at least some regularly reduced forms that are not very familiar to the listener are stored in episodic memory until abstraction rules can be generated. In contrast, observing that learning about reductions generalizes across talkers would be hard to explain based solely on storage of exemplars of the exposure talker's words and would therefore be a strong test of abstraction.

With evidence for both storage and abstraction processes, the results of this thesis argue against extreme models that allow only one of these two mechanisms to underlie the recognition of reduced words. Among the various hybrid models, those that do not include a mechanism for generalization (e.g., Ranbom & Connine, 2007) cannot account for the findings reported in Chapters 3 and 4. Furthermore, models that include generalization mechanisms must provide prelexical representations that are larger than a segment to be able to explain the adaptation to *ver*-reductions found in Chapter 3. The results of Chapter 4 suggest that these larger-than-segments units have to be fixed sequences of segments that occur in many words. It is thus likely that only morphemes satisfy this constraint.

The many ways that listeners adapt to reductions

This thesis has shown that listeners can adapt to reduced words so that they are able to recognize the same or other reduced words better. Which words are subsequently best recognized depends on how consistently the words that listeners have encountered had been reduced. If they were reduced in a highly consistent pattern, listeners could handle similar forms of reduction very well. If they were reduced in a very variable manner, listeners tended to tolerate very different forms of reduction. They were then flexible enough to adjust to reduced speech in general.

Flexible perceptual adjustments may provide weaker effects than learning about a consistent reduction pattern (see Chapter 4), but they are applicable to a broader range of reduced words. Listeners are thus well able to handle reductions and thus to cope with a phenomenon that is very frequent in everyday speech.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(4), 419-439.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language, 60*(1), 92-111.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer (Version 5.1.26). Amsterdam: Institute of Phonetic Sciences.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*(2), 707-729.

Brouwer, S. (2010). *Processing strongly reduced forms in casual speech.* Ph.D. thesis, Radboud University, Nijmegen.

Brouwer, S., Mitterer, H., & Huettig, F. (2012). Speech reductions change the dynamics of competition during spoken word recognition. *Language and Cognitive Processes, 27*(4), 539-571.

Brouwer, S., Mitterer, H., & Huettig, F. (2013). Discourse context and the recognition of reduced and canonical spoken words. *Applied Psycholinguistics, 34*, 519-539.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology, 58*(5), 412-424.

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

Cho, T., McQueen, J. M., & Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics, 35*(2), 210-243.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America, 116*(6), 3647-3658.

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*(3), 403-411.

Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience, 6*(3), 256-266.

Crone, P., & Fuchs, F. (2013, February 14). Scherzerl gmacht, keiner glacht. *Süddeutsche Zeitung*. Retrieved from http://www.sueddeutsche.de/muenchen/vorwuerfe-gegen-ex-arrichef-arnold-scherzerl-gmacht-keiner-glacht-1.1599819

Cutler, A., Eisner, F., McQueen, J. M., & Norris, D. (2010). How abstract phonemic categories are necessary for coping with speaker-related variation. In C. Fougeron, B. Kühnert, D. I. M. & N. Vallée (Eds.), *Laboratory Phonology 10*. Berlin: de Gruyter.

Dahan, D., Drucker, S. J., & Scarborough, R. A. (2008). Talker adaptation in speech perception: Adjusting the signal or the representations? *Cognition, 108*(3), 710-718.

Dodokay (Writer). (2010a). Neue Technik für die Polizei - Landesschau BW, *Die Welt auf Schwäbisch*. Germany: SWR.

Dodokay (Writer). (2010b). Promi-Talk mit Karl Lagerfeld, *Die Welt auf Schwäbisch*. Germany: SWR.

Dodokay (Writer). (2010c). Was ist lommelich? - Landesschau Baden-Württemberg, *Die Welt auf Schwäbisch*. Germany: SWR.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance, 25*(6), 1568-1578.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics, 67*(2), 224-238.

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time (L). *Journal of the Acoustical Society of America, 119*(4), 1950-1953.

Eisner, F., Melinger, A., & Weber, A. (2013). Constraints on the transfer of perceptual learning in accented speech. *Frontiers in Psychology, 4*.

Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch. A corpus-based study of the phonology-phonetics interface.* Ph.D. thesis, LOT, Utrecht.

Ernestus, M. (2009). The roles of reconstruction and lexical storage in the comprehension of regular pronunciation variants. *Proceedings of the 10th Annual Conference of the International Speech Communication Association (Interspeech 2009)* (pp. 1875-1878): Causal Productions Pty Ltd.

Ernestus, M., Baayen, H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language, 81*(1-3), 162-173.

Ernestus, M., & Warner, N. (2011a). An introduction to reduced pronunciation variants. *Journal of Phonetics, 39*(3), 253-260.

Ernestus, M., & Warner, N. (2011b). Speech reduction. [Special issue]. *Journal of Phonetics, 39*(SI).

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1276-1293.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*(2), 105-132.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12*(5-6), 613-656.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. [Review]. *Journal of Experimental Psychology: Learning Memory and Cognition, 22*(5), 1166-1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*(2), 251-279.

Gow, D. W. (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance, 28*(1), 163-179.

Gow, D. W. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics, 65*(4), 575-590.

Hanique, I., Ernestus, M., & Schuppler, B. (2013). Informal speech processes can be categorical in nature, even if they affect many different words. *Journal of the Acoustical Society of America, 133*(3), 1644-1655.

Hanulíková, A., & Weber, A. (2012). Sink positive: Linguistic experience with th substitutions influences nonnative word recognition. *Attention Perception & Psychophysics, 74*(3), 613-629.

Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics, 31*(3-4), 373-405.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*(4), 411-428.

IPdS. (1994). The Kiel Corpus of Spontaneous Speech. Kiel, Germany: Universität Kiel.

Jesse, A., & McQueen, J. M. (2011). Positional effects in the lexical retuning of speech perception. *Psychonomic Bulletin & Review, 18*(5), 943-950.

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

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Spontaneous Speech: Data and Analysis. Proceedings of the 1st Session of the 10th International Symposium* (pp. 29-54). Tokyo, Japan: The National International Institute for Japanese Language.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229-254). Amsterdam: John Benjamins.

Kemps, R., Ernestus, M., Schreuder, R., & Baayen, H. (2004). Processing reduced word forms: The suffix restoration effect. *Brain and Language, 90*(1-3), 117-127.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42*(3), 643-650.

Keune, K., Ernestus, M., van Hout, R., & Baayen, H. R. (2005). Social, geographical, and register variation in Dutch: From written MOGELIJK to spoken MOK. *Corpus Linguistics and Linguistic Theory, 1*, 183-223.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51*(2), 141-178.

Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language, 56*(1), 1-15.

Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008). First impressions and last resorts - How listeners adjust to speaker variability. *Psychological Science, 19*(4), 332-338.

Kuhn, D. (2005). Dodokay. Die Welt auf Schwäbisch. 2013, from http://www.dodokay.com/

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 637-675). Berlin: Mouton de Gruyter.

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics, 38*(1), 44-59.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Manuel, S. Y. (1992). Recovery of ‘‘deleted’’ schwa. In O. Engstrand & C. Kylander (Eds.), *Perilus: Papers from the Symposium on current phonetic research paradigms for speech motor control* (Vol. XIV, pp. 115–118): Stockholm: University of Stockholm.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*(1-2), 71-102.

Mattys, S. L., Bradlow, A. R., Davis, M. H., & Scott, S. K. (2012). Speech recognition in adverse conditions. [Special Issue]. *Language and Cognitive Processes, 27*(7-8).

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes, 27*(7-8), 953-978.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science, 32*(3), 543-562.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1-86.

McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 539-553.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*(6), 1113-1126.

McQueen, J. M., & Huettig, F. (2012). Changing only the probability that spoken words will be distorted changes how they are recognized. *Journal of the Acoustical Society of America, 131*(1), 509-517.

McQueen, J. M., Norris, D., & Cutler, A. (2006). The dynamic nature of speech perception. *Language and Speech, 49*, 101-112.

McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology, 60*(5), 661-671.

Mitterer, H. (2012). Ecological aspects of speech perception. [Research topic / Special Issue]. *Frontiers in Cognition*.

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

Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological abstraction in processing lexical-tone variation: Evidence from a learning paradigm. *Cognitive Science, 35*(1), 184-197.

Mitterer, H., Csépe, V., & Blomert, L. (2006). The role of perceptual integration in the recognition of assimilated word forms. *Quarterly Journal of Experimental Psychology, 59*(8), 1395-1424.

Mitterer, H., Csépe, V., Honbolygo, F., & Blomert, L. (2006). The recognition of phonologically assimilated words does not depend on specific language experience. *Cognitive Science, 30*(3), 451-479.

Mitterer, H., & Ernestus, M. (2006). Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics, 34*(1), 73-103.

Mitterer, H., & McQueen, J. M. (2009a). Foreign subtitles help but native-language subtitles harm foreign speech perception. *Plos One, 4*(11), A146-A150.

Mitterer, H., & McQueen, J. M. (2009b). Processing reduced word-forms in speech perception using probabilistic knowledge about speech production. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 244-263.

Mitterer, H., & Reinisch, E. (in press). No delays in application of perceptual learning in speech recognition: Evidence from eye tracking. *Journal of Memory and Language*.

Mitterer, H., & Russell, K. (2013). How phonological reductions sometimes help the listener. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 977-984.

Newman, R. S., & Sawusch, J. R. (2009). Perceptual normalization for speaking rate III: Effects of the rate of one voice on perception of another. *Journal of Phonetics, 37*(1), 46-65.

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review, 115*(2), 357-395.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*(2), 204-238.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60*(3), 355-376.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5*(1), 42-46.

Oostdijk, N. (2000). The Spoken Dutch Corpus Project. *ELRA Newsletter, 5*, 4-8.

Patterson, D., LoCasto, P. C., & Connine, C. M. (2003). Corpora analyses of frequency of schwa deletion in conversational American English. *Phonetica, 60*(1), 45-69.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*, 329-392.

Pitt, M. A. (2009). The strength and time course of lexical activation of pronunciation variants. *Journal of Experimental Psychology: Human Perception and Performance, 35*(3), 896-910.

Pitt, M. A., Dilley, L., & Tat, M. (2011). Exploring the role of exposure frequency in recognizing pronunciation variants. *Journal of Phonetics, 39*(3), 304-311.

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language, 39*(3), 347-370.

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America, 118*(4), 2561-2569.

Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language, 57*(2), 273-298.

Reinisch, E., Jesse, A., & McQueen, J. M. (2011). Speaking rate from proximal and distal contexts is used during word segmentation. *Journal of Experimental Psychology: Human Perception and Performance, 37*(3), 978-996.

Reinisch, E., & Sjerps, M. J. (2013). The uptake of spectral and temporal cues in vowel perception is rapidly influenced by context. *Journal of Phonetics, 41*(2), 101-116.

Reinisch, E., & Weber, A. (2012). Adapting to suprasegmental lexical stress errors in foreign-accented speech. *Journal of the Acoustical Society of America, 132*(2), 1165-1176.

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention Perception & Psychophysics, 71*(6), 1207-1218.

Schuppler, B., Ernestus, M., Scharenborg, O., & Boves, L. (2011). Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental transcriptions. *Journal of Phonetics, 39*(1), 96-109.

Shockey, L. (2003). *Sound Patterns of Spoken English*. Cambridge: Blackwell.

Sjerps, M. J., & McQueen, J. M. (2010). The bounds on flexibility in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 36*(1), 195-211.

Sjerps, M. J., Mitterer, H., & McQueen, J. M. (2011). Listening to different speakers: On the time-course of perceptual compensation for vocal-tract characteristics. *Neuropsychologia, 49*(14), 3831-3846.

Spinelli, E., & Gros-Balthazard, F. (2007). Phonotactic constraints help to overcome effects of schwa deletion in French. *Cognition, 104*(2), 397-406.

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: Academic.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*(4), 487-501.

Taft, M., Hambly, G., & Kinoshita, S. (1986). Visual and auditory recognition of prefixed words. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 38*(3), 351-365.

Trask, R. L. (2000). *The dictionary of historical and comparative linguistics*. Edinburgh: Edinburgh University Press.

Van de Ven, M., Tucker, B. V., & Ernestus, M. (2011). Semantic context effects in the comprehension of reduced pronunciation variants. *Memory & Cognition, 39*(7), 1301-1316.

Van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience, 13*(7), 967-985.

Van Son, R. J. J. H., & Pols, L. C. W. (1990). Formant frequencies of Dutch vowels in a text, read at normal and fast rate. *Journal of the Acoustical Society of America, 88*(4), 1683-1693.

Van Son, R. J. J. H., & Pols, L. C. W. (1992). Formant movements of Dutch vowels in a text, read at normal and fast rate. *Journal of the Acoustical Society of America, 92*(1), 121-127.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*(3), 374-408.

Warner, N., & Tucker, B. V. (2011). Phonetic variability of stops and flaps in spontaneous and careful speech. *Journal of the Acoustical Society of America, 130*(3), 1606-1617.

Witteman, M. J. (2013). *Lexical processing of foreign-accented speech: Rapid and flexible adaptation.* Ph.D. thesis, Radboud University, Nijmegen.

Witteman, M. J., Bardhan, N. P., Weber, A., & McQueen, J. M. (in press). Automaticity and stability of adaptation to a foreign-accented speaker. *Language and Speech*.

Witteman, M. J., Weber, A., & McQueen, J. M. (2013). Foreign accent strength and listener familiarity with an accent codetermine speed of perceptual adaptation. *Attention Perception & Psychophysics, 75*(3), 537-556.

Witteman, M. J., Weber, A., & McQueen, J. M. (in press). Tolerance for inconsistency in foreign-accented speech. *Psychonomic Bulletin & Review*.

Samenvatting en conclusies

Dit proefschrift heeft twee lijnen van onderzoek gecombineerd die tot dusver onafhankelijk van elkaar waren, maar die zich beide bezig houden met hoe luisteraars ongunstige luisteromstandigheden bij het herkennen van gesproken spraak overwinnen: onderzoek naar fonologische reducties en naar sprekeradaptatie. Aangezien sprekers verschillen in de mate waarin ze fonologische reducties produceren, zou het nuttig zijn voor luisteraars om zich aan de "reductiestijl" van individuele sprekers aan te passen. Dit was de reden om de twee hierboven genoemde onderzoekslijnen te combineren. Het kijken naar of en hoe luisteraars zich aan sprekerspecifieke reducties aanpassen heeft ook implicaties voor het debat over de structuur van het mentale lexicon. Het zou kunnen dat luisteraars de meeste gereduceerde woorden die ze tegenkomen in een episodisch lexicon opslaan. Men zou dan duidelijke aanwijzingen voor woordspecifiek leren moeten vinden. Als luisteraars daarentegen voornamelijk over het spraaksignaal abstraheren om gereduceerde woorden te herkennen in een fonologisch abstract lexicon, zou men moeten kunnen waarnemen dat het leerproces wordt gegeneraliseerd naar nieuwe woorden die op dezelfde wijze gereduceerd zijn. Als luisteraars kunnen abstraheren over de reductiepatronen die ze tegenkomen, rijst bovendien de vraag hoeveel verval van het spraaksignaal ze aankunnen. Dit proefschrift stelde dus de vraag of luisteraars uitsluitend overweg kunnen met segmentele reducties of ook met reducties van grotere linguïstische eenheden zoals de lettergreep. Daarnaast behandelt dit proefschrift het tijdsverloop van adaptatie aan reducties. Ik heb naar twee aspecten van het tijdsverloop gekeken: ten eerste naar het tijdstip waarop het leren wordt toegepast bij de herkenning van een gereduceerd woord en ten tweede naar hoeveel ervaring ervoor nodig is om je aan een specifiek reductiepatroon aan te passen.

Samenvatting van de resultaten

Onderzoek naar de hoeveelheid blootstelling die nodig is om een leerproces in gang te zetten en onderzoek naar de onderliggende leerfunctie zou mogelijke verschillen aan het licht kunnen brengen tussen verschillende vormen van adaptatie. Het kan zijn dat het leerproces een gestaag proces is dat op een gegeven moment een plateau bereikt of dat het op een abruptere manier gebeurt (d.w.z. „nu snap ik het“). Maar zelfs voor het veelvuldig onderzochte leerproces van sprekereigenaardigheden (e.g., McQueen, Cutler, & Norris, 2006; Norris, McQueen, & Cutler, 2003) was er geen eerder werk dat duidelijk aangaf wanneer en met welk tijdsverloop het leerproces zich voltrok. Daarom was een van de eerste stappen van dit project om m.b.v. een eyetracker het tijdsverloop van het perceptuele leren over ambigue fricatieven te onderzoeken. Dit leerproces zou overeenkomsten kunnen vertonen met het proces van aanpassing aan segmentele reducties, waardoor dit onderzoek nuttige informatie voor verder onderzoek over reducties zou kunnen opleveren.

Het doel van het experiment in Hoofdstuk 2 was om te bepalen hoeveel exemplaren een luisteraar moet horen om ervan te leren en hoe de leerfunctie eruit ziet. Daarom werden testtrials tussengevoegd tussen trainings- en fillertrials. In de trainingstrials werden twee groepen Nederlandse proefpersonen blootgesteld aan woorden die de klank $[^{s}/_{f}]$ bevatten. De proefpersonen leerden om deze ambigue klank ofwel als /s/ (in de /s/-tendens groep) ofwel als /f/ (in de /f/-tendens groep) te interpreteren op basis van lexicale tendensen. Dat wil zeggen dat ze deze vreemde klank in woorden hoorden die of op /s/ of op /f/ eindigen en dat ze op de orthografische vorm van dit woord moesten klikken (bv. op *radijs* of op *olijf*). In de testtrials hoorden de proefpersonen de ambigue klank aan het eind van een woord dat zowel op /s/ als op /f/ kon eindigen (bv. gesproken *doo*$[^{s}/_{f}]$ kon refereren aan zowel geschreven *doos* als aan geschreven *doof*). Leren trad op na blootstelling aan 10 trainingsitems en 10 tijdelijke minimale paren (d.w.z. woorden zoals *gister* en *giftig,* die nog ambigu waren toen luisteraars de ambigue klank hoorden die de fricatief verving). Ik nam dus resultaten waar die vergelijkbaar zijn met die van Kraljic en Samuel (2007), waar leren ook optrad na blootstelling aan 10 trainingsitems. Zodra het leereffect zich had voorgedaan, bleef het stabiel gedurende het verdere experiment; het werd niet sterker met extra training. Leren lijkt dus op een abrupte manier op te treden. Het is interessant om op te merken dat het leereffect zich alleen laat binnen een trial voordeed (ca. 900 ms na begin van het doelwoord). Mitterer en Reinisch (in press) hebben verder onderzoek gedaan naar dit verschijnsel, in een eyetracking studie met een geblokt blootstelling-test design. Zij ontdekten dat perceptueel leren effect heeft zodra dit mogelijk is, d.w.z. 200 ms na het begin van de fricatief. Het is dus waarschijnlijk dat de tussengevoegde testtrials (vooral de vroege) het leren nadelig beïnvloed hebben.

Omdat het design waarbij testtrials ingevoegd waren tussen trainingstrials mogelijk problematisch bleek te zijn voor de waarneming van leereffecten, is in Hoofdstuk 3 en 4 een klassiek design gebruikt waar blootstellings- en testtrials in blokken verdeeld waren. Ik heb niet meer geprobeerd te achterhalen hoeveel items er nodig zijn om te leren. In plaats daarvan heb ik me geconcentreerd op het perceptueel leren over reducties.

Het doel van Hoofdstuk 3 was te testen of luisteraars zich kunnen aanpassen aan reducties van segmenten en lettergrepen in doorlopende spraak. De focus van dit hoofdstuk lag hierbij op de generalisatie van het leren. Twee groepen Nederlandse proefpersonen werden blootgesteld aan verschillende soorten reducties en er werd getest of dit hen hielp bij het herkennen van andere woorden die op een soortgelijke wijze waren gereduceerd. Een dergelijke generalisatie zou laten zien dat luisteraars prelexicale abstractie gebruiken.

De twee reductiesoorten die zijn onderzocht, zijn de verzwakking van een /b/ aan het begin van een woord tot een labiodentale approximant ([ʋ]) en de verkorting van het voorvoegsel *ver*- tot een lange [fː]. Deze twee soorten reducties verschillen in de grootte van de gereduceerde eenheid (segment versus lettergreep), zodat adaptatie ons meer informatie zou kunnen verschaffen over de grootte(n) van de abstracte eenheden die op het prelexicale niveau gerepresenteerd zijn. Als leren voor beide soorten reducties optreedt, zou dat impliceren dat prelexicale abstractie gebruik kan maken van eenheden die groter zijn dan een enkel segment. Proefpersonen in de zogenoemde segmentele groep werden aan gereduceerde /b/-woorden blootgesteld, terwijl ze de woorden die met *ver*- begonnen in canonieke vorm hoorden. De syllabische reductiegroep werd blootgesteld aan gereduceerde *ver*-woorden, terwijl de woorden die met /b/ begonnen in canonieke vorm voorkwamen. Op deze manier fungeerden de twee groepen als controlegroep voor elkaar. In de testfase hoorden beide groepen beide soorten reducties in nieuwe woorden en ik heb vergeleken hoe goed ze deze gereduceerde woorden herkenden.

In Experiment 3.1 werd een leereffect voor de syllabische reductiegroep waargenomen, maar niet voor de segmentele reductiegroep. Dat wil zeggen dat de syllabische reductiegroep bij *ver*-reducties een grotere voorkeur voor het doelwoord had dan de segmentele reductiegroep, maar dat de segmentele reductiegroep bij /b/-reducties geen grotere voorkeur voor het doelwoord had dan de syllabische reductiegroep. Dit nuleffect zou toegeschreven kunnen worden aan verschillende aspecten van de generering van de stimuli. Het aspect dat me het meest waarschijnlijk leek, was een ongeschikte reductiecontext: De spreker sprak de schwa die aan het gereduceerde doelwoord vooraf ging stemloos uit, met als gevolg dat de /b/ niet in een intervocale positie optrad. Deze intervocale context zou echter cruciaal kunnen zijn voor het optreden van /b/-reducties. Daarom is de draagzin in een volgend experiment veranderd. In Experiment 3.2 werden vervolgens leereffecten voor zowel de syllabische reductiegroep als de segmentele reductiegroep gevonden, wat erop wijst dat de fonetische context belangrijk is om te bepalen hoe luisteraars gereduceerde vormen verwerken.

Net zoals in Hoofdstuk 2 komen de leereffecten laat in de eyetrackingsgegevens naar boven. De voorkeur die de proefpersonen hadden voor het gereduceerde doelwoord was pas vanaf ongeveer 600-800 ms na het begin van het doelwoord zichtbaar.

Omdat de gereduceerde testwoorden niet in de blootstellingsfase voorkwamen, werd het leren over /b/- en *ver*-reducties gegeneraliseerd over woorden. Als leren op een lexicaal of op een postlexicaal niveau plaatsvindt, zou er geen generalisatie van leren waargenomen worden. Maar als leren op een prelexicaal niveau plaatsvindt, zou het automatisch voor alle woorden met hetzelfde reductiepatroon van toepassing moeten zijn. Het feit dat leren over sublexicale eenheden ook toegepast werd op andere woorden suggereert dat de mechanismen die compenseren voor segmentele en syllabische reducties zich op een prelexicaal niveau bevinden. De resultaten suggereren dus dat er een abstracte verwerkingsmodus betrokken is bij de lexicale toegang. Toch tegenwoordige abstracte modellen moeten herzien worden, omdat ze geen representaties die groter zijn dan een segment op de prelexicale niveau bevatten (die nodig zijn om adaptatie aan *ver*-reducties te kunnen motiveren). Het moet echter opgemerkt worden dat de syllabische reductie die in Hoofdstuk 3 onderzocht werd tegelijkertijd ook de reductie van een morfeem was. Het is dus niet uit te sluiten dat representaties op prelexicaal niveau die groter zijn dan één segment uitsluitend uit morfemen bestaan.

Bovendien waren de leereffecten die in Hoofdstuk 3 zijn gevonden niet sterk. Hiervoor zijn twee mogelijke verklaringen. Enerzijds zou het zo kunnen zijn dat luisteraars hun kennis over reducties niet alleen toepassen op de reductiesoort waar ze bekend mee zijn, maar ook op andere soorten reducties. In dat geval kunnen beide groepen luisteraars heel goed omgaan met /b/- en *ver*-reducties, waarbij de bekendere reductiesoort een klein voordeel heeft. Anderzijds, en dit is het tegengestelde scenario, zou het zo kunnen zijn dat leren over reducties overwegend woordspecifiek is. Beide groepen luisteraars gaan dan vrij slecht om met /b/- en *ver*-reducties, omdat er alleen minimale generalisatie binnen en tussen reductiesoorten is.

Hoofdstuk 4 behandelt deze kwestie. Om na te gaan of leren over reductiesoorten gegeneraliseerd wordt, werd een controlegroep geïntroduceerd die helemaal geen reducties hoorde. Daarnaast waren er twee experimentele groepen, die, net als in Hoofdstuk 3, aan segmentele danwel aan syllabische reducties blootgesteld werden. Als er sprake is van generalisatie, zou ik verwachten dat de controlegroep bij beide soorten reducties overtreffen wordt door alle twee de experimentele groepen. Om de kwestie van woordspecificiteit aan te pakken werden woorden in de blootstellings- en de testfase herhaald. Als leren voornamelijk woordspecifiek blijkt te zijn, zouden de leereffecten veel groter zijn wanneer woorden herhaald worden dan wanneer ze niet herhaald worden.

In Experiment 4.1 en 4.2 kwamen de gereduceerde woorden uit de testfase eerder al voor in de blootstellingsfase. De segmentele reductiegroep werd blootgesteld aan woorden die met een /b/ beginnen, maar waar de beginmedeklinker tot een bilabiale nasaal gereduceerd werd. De syllabische reductiegroep werd blootgesteld aan woorden die met CVC beginnen, maar waaruit de onbeklemtoonde volle klinker werd verwijderd. Er is gekozen voor de deletie van volle klinkers om verder te onderzoeken of de adaptatieprocessen beïnvloed worden door de morfemische status van de lettergrepen die bij reductieprocessen betrokken zijn. De syllabische reducties die in Hoofdstuk 3 onderzocht werden, waren reducties van een voorvoegsel. De canonieke vormen van deze reducties waren daarom uitermate constant wat hun segmentele structuur betreft (d.w.z. hetzelfde foneem kwam altijd in een bepaalde positie voor). Hoofdstuk 4 testte of luisteraars zich ook kunnen aanpassen aan de reductie van niet-morfemische lettergrepen die noodzakelijkerwijs een grotere variatie in hun segmentele structuur vertonen dan voorvoegsels. De prestaties van de twee experimentele groepen werden beide vergeleken met de prestatie van de controlegroep, en niet met elkaar. Zowel voor de segmentele als voor de syllabische reductiegroep werden leereffecten waargenomen voor herhaalde woorden. Om te testen of deze effecten echt woordspecifiek waren en niet door abstractieprocessen veroorzaakt werden, is vervolgens Experiment 4.3 uitgevoerd. Net als in Hoofdstuk 3 lag de focus bij dit experiment op de generalisatie van leren. Leren over klinkerdeleties werd niet gegeneraliseerd naar andere woorden waar een klinker uit werd verwijderd. De betreffende leereffecten uit Experiment 4.1 en 4.2 waren dus inderdaad woordspecifiek. Leren over /b/-reducties werd echter wel gegeneraliseerd naar nieuwe gereduceerde /b/-woorden. Het betreffende leereffect dat in Experiment 4.1 werd gevonden was dus niet woordspecifiek. Er is een belangrijk verschil tussen de twee reductiesoorten dat deze verschillende soorten van leerprocessen zou kunnen verklaren. Dit verschil betreft de constantheid van de input. Waar de klinkerdeleties een heel gevarieerd reductiepatroon lieten zien, was het patroon van de /b/-reducties zeer uniform. Het ziet er dus naar uit dat een uniform reductiepatroon abstractieprocessen vergemakkelijkt, terwijl gevarieerde reductiepatronen verwerkt worden door het opslaan van individuele exemplaren van gereduceerde woorden.

Er werd uitsluitend een generalisatie van leren over reductiesoorten gevonden voor de syllabische reductiegroep. Proefpersonen in deze groep konden hun kennis van klinkerdeleties gebruiken om gereduceerde /b/-woorden beter te herkennen. Voor de generalisatie van het leren over /b/-reducties naar klinkerdeleties werden daarentegen inhiberende effecten waargenomen (d.w.z. dat de segmentele reductiegroep een kleinere voorkeur vertoonde voor doelwoorden waar een klinker uit verwijderd was dan de controlegroep). De meest waarschijnlijke verklaring voor deze asymmetrische bevinding moet wederom gezocht werden in de heterogeniteit van het reductiepatroon. Luisteraars die aan een variabel reductiepatroon blootgesteld worden, lijken hierdoor toleranter te worden voor alle vormen van reductie.

De faciliterende effecten voor woordspecifiek leren en voor het type leren dat binnen een reductiesoort gegeneraliseerd wordt, traden laat in de trial op (ongeveer 1100 ms na het begin van het doelwoord). De inhiberende effecten en de faciliterende effecten voor generalisatie van het leren over reductiesoorten daarentegen traden vroeg op (vanaf 200 ms, zoals in Experiment 4.2).

Dat er verschillende soorten leereffecten gevonden zijn, suggereert dat luisteraars verschillende middelen gebruiken om gereduceerde woorden te herkennen: Ze slaan de reducties op die ze zijn tegengekomen (het woordspecifieke leereffect), en ze bouwen abstracte regels zodat ze hun kennis indien mogelijk op nieuwe gereduceerde woorden kunnen toepassen (gegeneraliseerd leren binnen een reductiesoort). Gegeneraliseerd leren lijkt afhankelijk te zijn van de uniformiteit van de gereduceerde input. Bij de reducties van niet-morfemische lettergrepen was de input te variabel om erover te abstraheren. In plaats daarvan bleek het perceptuele systeem flexibeler te zijn in het tolereren van andere vormen van reductie (gegeneraliseerd leren over reductiesoorten).

Hoofdstuk 5 keerde terug naar de vraag hoeveel ervaring er nodig is voor adaptatie. In Hoofdstuk 2 werd al op deze vraag ingegaan, maar er werd een andere tijdschaal onderzocht. In plaats van de vraag te stellen hoeveel woorden een luisteraar moet tegenkomen om zich te kunnen aanpassen, ging Hoofdstuk 5 in op die verschillen in verwerking, die door een verschillende mate van ervaring (van een paar weken tot meerdere jaren) met een bepaald patroon veroorzaakt worden. Hiervoor werd een quasi-experimentele benadering gebruikt, waarin ervaring niet experimenteel gemanipuleerd werd, maar waarin groepen proefpersonen met verschillende hoeveelheden ervaring getest werden. Bovendien werd korte-termijn adaptatie opnieuw onderzocht. Er werd onderzocht of luisteraars, vooral degenen in de minder ervaren groep, in de loop van het experiment beter worden in het herkennen van gereduceerde woorden.

Met behulp van ERP's werd de adaptatie aan twee soorten Duitse voorvoegselreducties op de lange en de korte termijn onderzocht. De afwijking tussen de gereduceerde woordvormen en hun canonieke uitspraak zou een fonologische mismatch-detectie kunnen veroorzaken en daardoor een N200 effect kunnen oproepen. De afwijking zou echter ook moeilijkheden kunnen veroorzaken bij de toegang tot het mentale lexicon, en ook in een later stadium wanneer het woord in het zinsverband moet worden geïntegreerd, waardoor een N400 zou kunnen ontstaan. Hoe groter deze effecten zijn, hoe meer moeilijkheden luisteraars dus ondervinden bij het herkennen van gereduceerde vormen. Een van de twee onderzochte reductiesoorten (de reductie van *ge*- tot [g]/[k]) komt veelvuldig voor in Zuid-Duitsland en Oostenrijk, terwijl de andere soort reductie (de reductie van *ver*- tot [f]) relatief zeldzaam is in alle delen van Duitsland en Oostenrijk. Een Noord-Duitse groep, een migrantengroep (bestaande uit proefpersonen die van Noord- naar Zuid-Duitsland verhuisd waren) en een Zuid-Duitse groep luisterden naar zinnen die deze reducties bevatten, terwijl een controlegroep naar dezelfde zinnen luisterde, maar hierin slechts ongereduceerde vormen hoorde. Beide reductiesoorten leidden in de helft van de gevallen tot illegale fonotactische sequenties. Reducties die leiden tot illegale fonotactische sequenties, kunnen vroege verwerkingsmoeilijkheden veroorzaken, maar deze moeilijkheden werden snel overwonnen — dit in tegenstelling tot reducties die legale sequenties tot gevolg hebben. Deze reducties resulteerden in aanhoudende verwerkingskosten. De hogere verwerkingskosten voor "legale" reducties zijn waarschijnlijk te wijten aan de toegenomen lexicale competitie. Er worden immers niet alleen competitors geactiveerd die met hetzelfde voorvoegsel beginnen, maar daarnaast worden er eveneens competitors geactiveerd die met hetzelfde consonantcluster beginnen.

Voor *ge*-reducties lieten de resultaten zowel een voordeel van langdurige blootstelling zien voor de Zuid-Duitse luisteraars als snelle perceptuele aanpassingen voor de Noord-Duitse en migrantenluisteraars die op zijn minst enigszins vertrouwd waren met *ge*-reducties. Voor *ver*-reducties toonden de Zuid-Duitsers perceptuele flexibiliteit, terwijl de Noord-Duitsers en vooral de migranten aanhoudende verwerkingsmoeilijkheden hadden. Het gemak waarmee de Zuid-Duitsers de *ge*-reducties verwerkten, suggereert dat ze deze gereduceerde vormen in hun mentale lexicon opgeslagen hebben. De Zuid-Duitsers toonden ook minder verwerkingsmoeilijkheden voor *ver*-reducties vergeleken met de Noord-Duitse en de migrantengroep. Omdat alle drie de groepen relatief onbekend waren met *ver*-reducties, lijkt het waarschijnlijk dat de Zuid-Duitsers hun grotere ervaring met voorvoegselreducties in het algemeen hebben uitgebuit. Terwijl de Noord-Duitse en de migrantengroep vrij onbekend waren met *ver*-reducties, was hun beperkte ervaring met *ge*-reducties blijkbaar voldoende om hun perceptuele systeem flexibel aan te kunnen passen. Ze hadden moeilijkheden met de *ge*-reducties in het N200 tijdsinterval, maar deze kwamen ze snel te boven. In het N400 tijdsinterval verwerkten ze deze reducties dus zoals de Zuid-Duitse luisteraars dat deden. Vertrouwdheid met een bepaalde reductiesoort is dus vereist voor een moeiteloze herkenning van gereduceerde woorden.

In tegenstelling tot de eyetracking experimenten in Hoofdstuk 3 en 4 werd geen adaptatie op korte termijn (d.w.z. adaptatie in de loop van het experiment) gevonden. Er vonden echter al vroeg perceptuele aanpassingen plaats. Reducties die tot verwerkings-moeilijkheden leidden in het N200 tijdsvenster, leverden in het N400 tijdsvenster immers geen problemen meer op.

Conclusies

De experimentele hoofdstukken van dit proefschrift hebben inzicht verschaft in specifieke aspecten van de adaptatie aan segmentele en syllabische reducties. Er waren aanwijzingen gevonden dat er verschillende soorten mechanismen een rol spelen wanneer luisteraars zich aan gereduceerde spraak aanpassen. Dit heeft consequenties voor modellen die beschrijven hoe gesproken woorden worden herkend. De resultaten suggereren verder dat adaptatie aan reducties niet een geheel automatisch proces is, maar veeleer door bepaalde factoren beperkt wordt. Wat betreft het tijdsverloop lijkt adaptatie aan reducties te verschillen van perceptueel leren over sprekereigenaardigheden.

Processen die bij adaptatie aan reducties betrokken zijn

Er lijken verschillende leermechanismen betrokken te zijn bij de herkenning van gereduceerde woordvormen. Experiment 4.1 en 4.2 (over woordspecifiek leren) en Experiment 5 (dat geen verwerkingsmoeilijkheden voor zeer bekende vormen van reducties liet zien) suggereerden dat luisteraars gereduceerde uitspraakvarianten opslaan. Experiment 3.1, 3.2 en 4.3 gaven aan dat abstractieprocessen een belangrijke rol spelen bij het herkennen van woorden die op dezelfde manier gereduceerd zijn en die niet eerder gehoord waren. Tenslotte leverde de manier waarop luisteraars omgaan met woorden die op een onverwachte manier gereduceerd waren (de nieuwe reductiesoorten in de testfasen van Experiment 4.1 en 4.3 en de *ge*-reducties die voor de Noord-Duitse proefpersonen in Hoofdstuk 5 slechts beperkt bekend waren) een aanwijzing voor de flexibiliteit van het perceptuele systeem.

Deze verschillende mechanismen vullen elkaar in verschillende situaties mogelijk aan. Het opslaan van gereduceerde vormen is voordelig als men later hetzelfde woord, op dezelfde manier gereduceerd, nogmaals tegenkomt, vooral als men een bepaalde gereduceerde vorm vaak tegenkomt. De generalisatie van leren over reducties daarentegen helpt luisteraars om nieuwe gereduceerde woorden te herkennen die zij nog niet eerder waren tegengekomen. Maar generalisatieprocessen lijken beperkt te worden door factoren als de uniformiteit van de input (tenminste voor generalisatie binnen reductiesoorten). De resultaten van Hoofdstuk 4, en vooral de inhiberende effecten, suggereren dat generalisatie van een uniform reductiepatroon naar een variabele reductiepatroon niet mogelijk is. Dit laat echter de mogelijkheid open dat er een transfer zou kunnen optreden van het leren van één uniform patroon naar een ander uniform reductiepatroon. Dit lijkt plausibel als het om reductiesoorten gaat die erg op elkaar lijken. Toekomstig onderzoek zou moeten uitwijzen of er een sterk generalisatie-effect over reductiesoorten gevonden kan worden voor reducties die in grootte en soort op elkaar lijken (bv. schwa-reductie in de Duitse voorvoegsels *be*- en *ge*-, of reductieprocessen die een vergelijkbare invloed hebben op verschillende plosieven. Of leren dan ook gegeneraliseerd kan worden van een uniform reductiepatroon naar een ander uniform reductiepatroon dat aanzienlijk verschilt (bv. de /b/-reducties en de *ver*-reducties die in Hoofdstuk 3 onderzocht werden) blijft verder een open vraag.

Wanneer men nieuwe gereduceerde vormen tegenkomt die op een onverwachte manier gereduceerd zijn, had het perceptuele systeem flexibel genoeg kunnen zijn om zich daaraan aan te passen — mits het van tevoren voldoende input had gehad. Waarschijnlijk moet deze voorafgaande input variabel genoeg zijn, zodat het systeem elke vorm van reductie tolereert (bv. zoals de klinkerdeleties in Hoofdstuk 4; proefpersonen die hieraan blootgesteld waren, konden vervolgens vrij goed omgaan met /b/-reducties), en/of voldoende lijken op de onverwachte gereduceerde vormen. De relatief zeldzame reductie van het voorvoegsel *ver*- (meestal als [fɐ] uitgesproken) naar [f] die in Hoofdstuk 5 onderzocht werd, leek in zoverre op de reductie van het voorvoegsel *ge*- naar [g]/[k] dat in beide gevallen een klinker verwijderd was. Maar alleen de Zuid-Duitse luisteraars die geen moeite hadden met het herkennen van de *ge*-reducties wisten ook vrij goed om te gaan met deze *ver*-reducties. Voldoende ervaring met *ge*-reducties (en zeer waarschijnlijk ook met andere vormen van voorvoegselreducties) en de relatief grote gelijkenis tussen de twee reductiesoorten vergemakkelijkte waarschijnlijk deze perceptuele aanpassing.

Beperkingen van het adaptatieproces

De bovengenoemde adaptatieprocessen werden door verschillende factoren beperkt. Ten eerste, en zoals reeds besproken, speelt de input een cruciale rol bij het bepalen van hoe luisteraars zich aan gereduceerde vormen aanpassen. Ten tweede geeft het feit dat leren voor /b/-reducties in Experiment 3.1 niet gegeneraliseerd werd aan dat adaptatie aan bepaalde reducties alleen kan plaatsvinden indien de juiste articulatorische context (bv. intervocale positie) aanwezig is. Luisteraars zijn gevoelig voor dergelijke probabilistische beperkingen (Mitterer & McQueen, 2009b) en passen zich niet aan als niet aan deze voorwaarden wordt voldaan.

Verder toonde Hoofdstuk 5 dat bekendheid met een bepaalde reductiesoort nodig is voor een moeiteloze herkenning van gereduceerde vormen. In Hoofdstuk 3 en 4 is geverifieerd dat de gekozen reductiesoorten daadwerkelijk in natuurlijke spraak voorkomen. Proefpersonen waren deze soorten waarschijnlijk dus al eens tegengekomen en waren er op zijn minst enigszins vertrouwd mee. Het feit dat er geen adaptatie aan ongewone reductievormen werd gevonden (de *ver*-reducties in Hoofdstuk 5 voor de Noord-Duitse en de migrantengroep), suggereert dat er een minimumblootstelling nodig is voordat adaptatie kan plaatsvinden.

Bovendien lijkt het merkwaardig dat enkele leereffecten die in Experiment 4.1 waargenomen werden niet gerepliceerd konden worden in de latere experimenten van Hoofdstuk 4. Alhoewel er in Experiment 4.3 op zijn minst een trend van generalisatie van het leren van klinkerdeleties naar /b/-reducties zichtbaar was, zowel in de eyetrackingsgegevens als in de gegevens m.b.t. het percentage correcte responses, werden in Experiment 4.2 geen aanwijzingen gevonden voor een effect van herhaalde /b/-reducties. Gevoelsmatig zouden luisteraars het eigenlijk gemakkelijker moeten vinden om een gereduceerd woord te herkennen wanneer ze het een tweede keer tegenkomen, of ze woorden nu via opslag- of via abstractieprocessen herkennen. Het nuleffect voor de herhaling van gereduceerde /b/-woorden in Experiment 4.2 suggereert dus dat adaptatie aan reducties niet automatisch plaatsvindt. Dit staat in tegenstelling tot datgene wat werd beweerd voor perceptueel leren over sprekereigenaardigheden (McQueen, et al., 2006) en adaptatie aan spraak met een buitenlands accent (Witteman, Bardhan, Weber, & McQueen, in press). Een eigenschap van sprekereigenaardigheden (hier: de vreemde uitspraak van een fricatief), spraak met een buitenlands accent (hier: verkeerd uitgesproken Nederlandse klinkers) en syllabische reducties, die verzwakte segmenten mogelijk niet in dezelfde mate bezitten, is opvallendheid. Luisteraars zijn zich er waarschijnlijk van bewust dat er segmenten ontbreken (zoals dat bij syllabische reducties gebeurt) en dat sommige segmenten niet zo uitgesproken worden als een moedertaalspreker dat zou doen. Een ambigue fricatief die een gelispel zou representeren is ook zeer opvallend. Dit in tegenstelling tot een verzwakte segment, dat mogelijk niet voldoende afwijkt van de canonieke uitspraak om adaptatie te garanderen—tenzij andere factoren de opvallendheid verhogen.

Een dergelijke factor zou voorspelbaarheid kunnen zijn. Als luisteraars het volgende, en mogelijk gereduceerde, woord kunnen voorspellen, helpt dat hen waarschijnlijk om zich aan reducties aan te passen. Op soortgelijke wijze suggereerden Jesse en McQueen (2011) dat het voor leren noodzakelijk is dat lexicale kennis beschikbaar is op het moment dat een ambigue klank verwerkt wordt. Ze hadden namelijk geen perceptueel leereffect kunnen vinden voor ambigue fricatieven die voorkwamen aan het begin van een woord dat zonder zinsverband gepresenteerd werd. Wanneer lexicale kennis tijdens de blootstelling werd verstrekt door het weergeven van de orthografische vorm van een gereduceerd woord op het scherm, werd een (gegeneraliseerd) leereffect waargenomen (zie Experiment 3.1, 3.2, 4.3). De gedrukte woorden op het scherm helpen moedertaalsprekers met de adaptatie aan reducties. De manier waarop die gebeurt lijkt waarschijnlijk op de manier waarop buitenlandse ondertitels niet-moedertaalsprekers helpen wanneer deze een film in deze buitenlandse taal kijken (Mitterer & McQueen, 2009a). In Experiment 4.2 waren de semantisch volle blootstellingszinnen alleen echter niet voldoende om een leereffect voor de herhaalde /b/-reducties te triggeren. De reden hiervoor zou kunnen liggen in het hoge aantal laag-frequente woorden en de moeilijkheid om zeer voorspelbare zinnen voor deze woorden te vormen. In dagelijkse communicatie helpt de bredere discourse context bij het verhogen van de voorspelbaarheid van volgende woorden en het zou het vermogen van de proefpersonen om zich aan gereduceerde woorden aan te passen kunnen vergroten. Een toekomstig experiment zou deze hypothese kunnen testen door te onderzoeken of een duidelijk leereffect gevonden kan worden voor herhaalde hoog-frequente woorden die op een uniforme manier gereduceerd zijn, als ze in de blootstellingsfase in zeer voorspelbare zinnen of zelfs in een bredere discourse context voorkomen. Het lijkt dus plausibel dat het voordelig voor de leerproces is dat mogelijke lexicale kandidaten al geactiveerd worden voordat het gereduceerde signaal verwerkt wordt. Als men de segmentele reducties opvallender maakt door de voorspelbaarheid van hun canonieke vorm te verhogen, lokt dat waarschijnlijk voorspellingsfouten uit, wat op zijn beurt tot impliciet leren zou kunnen leiden (Clark, 2013; Pickering & Garrod, 2013).

Concluderend kan gesteld worden, dat er geen eenvoudig verhaal te vertellen is over wanneer adaptatie aan reducties kan slagen. Of luisteraars zich aan gereduceerde woordvormen kunnen aanpassen, hangt af van een samenspel van verschillende factoren, zoals uniformiteit van de input, geschikte articulatorische context, bekendheid, opvallendheid en voorspelbaarheid. Een andere mogelijke factor die het adaptatieproces zou kunnen beïnvloeden en die in dit proefschrift niet onderzocht werd, is hoe consequent de spreker reducties produceert. In de blootstellingsfasen van de experimenten in dit proefschrift werd een spreker nagebootst die alle woorden reduceert die voldoen aan de voorwaarden voor het optreden van een specifieke reductiesoort (d.w.z. woorden die onbeklemtoonde en niet-gereduceerde /b/s bevatten werden niet opgenomen in het experimentmateriaal van Hoofdstuk 3 en 4). Dit leidt vervolgens tot de vraag of luisteraars zich ook aan reducties kunnen aanpassen als ze een spreker bijvoorbeeld slechts af en toe /b/-reducties horen produceren. Witteman (2013) onderzocht een soortgelijke kwestie in Hoofdstuk 5 van haar proefschrift door te testen of moedertaalsprekers zich aan een (niet-moedertaal-)spreker kunnen aanpassen wiens uitspraak soms geaccentueerd klonk en verder klonk als die van een moedertaalspreker. Zij vond dat luisteraars zich inderdaad konden aanpassen, maar dat ze daarvoor meer tijd nodig hadden dan luisteraars die alleen aan geaccentueerde woorden werden blootgesteld.

Het tijdsverloop van adaptatie aan reducties

In dit proefschrift werd niet alleen onderzocht onder welke omstandigheden wat voor soort adaptatiemechanismen optreden, er is ook gekeken naar wanneer leren optreedt. Hoofdstuk 5 liet geen effecten van adaptatie op korte termijn zien, maar verschafte informatie over het tijdsverloop van de verwerking van gereduceerde vormen. Tenzij een soort reductie zeer frequent voorkomt (wat in Hoofdstuk 3 en 4 niet het geval was) lokken gereduceerde woorden een N200 effect uit. Een algemeen N400 effect werd alleen gevonden voor zeer ongewone vormen reducties bij luisteraars die er niet aan gewend zijn om met gereduceerde vormen om te gaan. Reducties die tot legale fonotactische sequenties leidden, lokten echter een negatievere N400 uit dan reducties die leidden tot illegale fonotactische sequenties. De eerstgenoemde veroorzaken dus meer langdurige verwerkingsmoeilijkheden en deze op een legale manier gereduceerde vormen werden in Hoofdstuk 3 en 4 met eyetracking paradigma's onderzocht.

De resultaten van deze eyetracking experimenten laten zien dat het moment waarop het leren over reducties zijn intrede doet afhankelijk kan zijn van de specificiteit van het leerproces. Faciliterende en inhiberende generalisatie-effecten over reductiesoorten die—in Hoofdstuk 4—waarschijnlijk niet specifiek zijn voor een bepaald segment of woord, werden vroeg waargenomen (vanaf 200 ms na het begin van het doelwoord). Dit in tegenstelling tot het woordspecifieke effect dat in Experiment 4.1 werd gevonden en dat vrij laat optrad (het begon pas bij 1100 ms na het begin van het doelwoord). De generalisatie-effecten binnen reductiesoorten die in Hoofdstuk 3 en 4 waargenomen werden, traden net zo laat op. Maar preferenties voor het concurrerende woord werden, als ze optraden, veel vroeger waargenomen, soms al 300 ms na het begin van het doelwoord. Dit lijkt te suggereren dat de bottom-up informatie van het spraaksignaal een sterke invloed op de spraakherkenning heeft (e.g., Marslen-Wilson, 1987) en dat leren tijd nodig heeft om de interpretatie te overwinnen die door het spraaksignaal bepaald wordt (d.w.z. "wat-je-hoort-is-wat-je-krijgt").

Dit lijkt plausibel voor woordspecifiek leren, vooral als het gereduceerde woord erg op een ander canoniek woord lijkt. Wanneer men de gereduceerde spraakinput *P'rijs* hoort, moet de activatie van *Parijs* op een bepaald moment zegevieren over de activatie van *prijs*. In het begin is de activatie van *prijs* waarschijnlijk sterker, omdat deze betekenis veel vaker voorkomt. Sprekerspecifieke informatie (bv. over de neiging van deze spreker om woorden zoals *Parijs* te reduceren) moet dan invloed uitoefenen en de perceptie ten gunste van de kandidaat *Parijs* verschuiven. Dit kan enige tijd duren (d.w.z. het hoeft niet meteen te gebeuren).

De effecten van op regels gebaseerd leren lijken echter niet noodzakelijkerwijs vertraagd te zijn. Mitterer en Reinisch (in press), die een follow-up studie voor Hoofdstuk 2 van dit proefschrift uitvoerden, vonden vroege leereffecten in hun eyetracking data. Ze concludeerden dat perceptueel leren inderdaad perceptueel en niet post-perceptueel is zodra het voltooid is. De late effecten van Hoofdstuk 2 kunnen dus toegeschreven worden aan de interferentie van de tussengevoegde testtrials. Maar hoe staat het met de late effecten in Hoofdstuk 3 en 4?

Mitterer en Reinisch (in press) gebruikten een eyetracking paradigma met een klassiek blootstelling-test-design en namen vroege leereffecten waar. Het lijkt dus onwaarschijnlijk dat de late effecten in Hoofdstuk 3 en 4 veroorzaakt worden door de combinatie van een leerparadigma met een eyetracking taak die geschreven woorden gebruikt. Verder moesten de proefpersonen in de studie van Mitterer en Reinisch (in press) bewuste metalinguïstische beoordelingen maken over de spraakklanken die ze hoorden. Ze moesten op geschreven woorden klikken waarvan de gesproken realisaties afweken van hun canonieke uitspraak. Omdat deze taak geen invloed had op de timing van de effecten in Mitterer en Reinisch (in press), lijkt het onwaarschijnlijk dat dit in de experimenten die in Hoofdstuk 3 en 4 gepresenteerd zijn wel het geval was.

Wanneer perceptueel leren in het algemeen vroeg moet optreden, en dat ook doet, en wanneer methodologische problemen uitgesloten kunnen worden als mogelijke oorzaak, wat zou dan de reden kunnen zijn voor de late leereffecten die in dit proefschrift waargenomen werden? Een punt waarin Hoofdstuk 3 en 4 van de studie van Mitterer en Reinisch (in press) en andere perceptuele leerstudies (bv., McQueen, et al., 2006) verschillen, betreft de regels die luisteraars moeten genereren of moeten aanpassen om afwijkende klank(en) correct te herkennen. In het klassieke perceptuele leerparadigma moeten luisteraars leren om een van tevoren onbekende ambigue klank in *een* van twee mogelijke foneemcategorieën op te nemen. Luisteraars moeten dus een nieuwe regel voor deze klank genereren. Idealiter wordt de klank in 50% van de gevallen in één categorie ingedeeld en in de overige 50% van de gevallen in één andere categorie. Met uitzondering van de vooraf geteste [$^{b}/_{ʋ}$] klank in Experiment 3.1 waren echter geen van de gereduceerde klanken die in Hoofdstuk 3 en 4 gebruikt werden echt ambigu. De natuurlijk voorkomende [$^{b}/_{ʋ}$]s die in Experiment 3.2 als vervangers voor /b/ gebruikt werden, klonken waarschijnlijk meer als /ʋ/ dan als /b/. De lange [fː]s die de *ver*-s in Experiment 3.1 en 3.2 vervingen, klonken duidelijk meer als /f/ dan als *ver*-. Hetzelfde geldt voor de natuurlijke exemplaren van [m] die de /b/s in Hoofdstuk 4 vervingen en ook voor de CVC-woorden in dat hoofdstuk, omdat hier de hele klinker verwijderd werd en dus alleen de medeklinkerclusters achterbleven.

In Hoofdstuk 3 en 4 leerden luisteraars dus dat een bepaalde spreker bijvoorbeeld de neiging heeft een /b/ als een [m] uit te spreken en dat een bestaande klank (bv. [m]) daarom voor deze spreker op *twee* categorieën (d.w.z., /m/ en /b/) geprojecteerd kan worden. Hierdoor zou hun perceptie van een [m] verschoven kunnen zijn. In plaats van een [m] voornamelijk als /m/ te beoordelen, zouden ze het nu in 80% van de gevallen als /m/ en in 20% van de gevallen als /b/ kunnen beoordelen. Bij dit type leren krijgt de canonieke vorm aanvankelijk nog de voorkeur vanwege de hypotheses die door het spraaksignaal worden bepaald. Pas als later binnenkomende segmenten de canonieke vorm uitsluiten, kan er een leereffect optreden. Zodra luisteraars aanwijzingen krijgen dat een bepaalde klank op meer dan één categorie geprojecteerd kan worden, heeft het leerproces waarschijnlijk meer tijd nodig om zijn intrede te doen.

Studies over spraak met een buitenlands accent hebben eveneens gekeken naar het toewijzen van een klank aan meerdere categorieën. Hanulíková en Weber (2012) onderzochten het effect van linguïstische ervaring op de perceptie van verschillende substituties (/t/, /s/, en /f/) van de Engelse klank /θ/ door Nederlandse en Duitse luisteraars. Terwijl Duitse luisteraars de voorkeur gaven aan de /s/-varianten, gaven Nederlandse luisteraars de voorkeur aan de /t/-varianten, wat overeenstemt met hun ervaring en hun eigen producties. Hoewel de /f/-varianten perceptueel het meest op de realisaties van /θ/ leken, gaven noch de Nederlandse noch de Duitse luisteraars er de voorkeur aan.

Sommige studies over het toewijzingsprobleem in spraak met een buitenlands accent behandelden ook adaptatie-effecten in moedertaalsprekers. Deze concentreerden zich echter meestal op het adaptatieproces zelf en niet op het tijdsverloop. Eisner, Melinger, en Weber (2013) onderzochten bijvoorbeeld perceptueel leren over stemloze plosieven in Engelse moedertaalsprekers. De proefpersonen hoorden een Nederlandse spreekster die stemhebbende /d/s aan het eind van woorden stemloos uitsprak (d.w.z. ze sprak /d/s als [t]s uit). Na de blootstelling aan deze varianten werden Brits Engelse luisteraars in een cross-modal priming taak geprimed door woorden die op een soortgelijke wijze verkeerd uitgesproken waren. Dat wil zeggen dat de luisteraars zich aan deze varianten aangepast hadden en de afwijkende (stemloze) plosief als de bedoelde stemhebbende plosief /d/ herkenden. Hoewel woordherkenning met cross-modal priming online gemeten wordt (d.w.z. terwijl de herkenning plaatsvindt), geeft deze methode geen inzicht in de fijne temporele structuur van het herkenningsproces. Hij kan alleen laten zien dat er adaptatie heeft plaatsgevonden, maar hij kan geen informatie verschaffen over wanneer precies die optreedt, d.w.z. of er adaptatie optreedt zodra de afwijkende klank verwerkt wordt of pas enige honderden milliseconden daarna.

Witteman et al. (in press) onderzochten adaptatie aan Nederlands met een Hebreeuws accent en gebruikten ook het cross-modal priming paradigma. De Hebreeuwse spreekster sprak Nederlandse woorden die de korte klinker [ɪ] bevatten correct uit, maar ze verkortte de Nederlandse klinker [i] eveneens tot [ɪ]. Na een korte blootstelling van 3,5 minuten aan dit accent, waarin de proefpersonen de taak hadden om op een foneem te letten dat niet door het accent beïnvloed was, werd een faciliterende priming voor de afwijkende woorden waargenomen bij de Nederlandse luisteraars. Er werd geconcludeerd dat adaptatie aan spraak met een buitenlands accent snel en automatisch gebeurt.

Terwijl onderzoek naar spraak met een buitenlands accent bevestigt dat luisteraars zich kunnen aanpassen aan klanken die aan meer dan een categorie toegewezen kunnen worden, presenteert dit proefschrift de eerste studies die ook naar het tijdsverloop van een dergelijk adaptatieproces kijken. In tegenstelling tot adaptatie aan gereduceerde spraak lijkt adaptatie aan spraak met een buitenlands accent automatisch te zijn, evenals perceptueel leren over sprekereigenaardigheden. Toekomstig onderzoek naar het tijdsverloop van adaptatie aan spraak met een buitenlands accent zou echter kunnen tonen dat leren, vanwege de toewijzingsproblemen, tijd nodig heeft om van kracht te worden. Het zou kunnen dat adaptatie aan spraak met een buitenlands accent in dit opzicht meer lijkt op adaptatie aan gereduceerde spraak dan op perceptueel leren over spreker-eigenaardigheden.

Modellen over de herkenning van gesproken woorden

Wat vertellen de verschillende vormen van adaptaties aan reducties die in dit proefschrift gevonden werden ons over hoe woorden in het mentale lexicon opgeslagen worden en, daarmee samenhangend, over hoe ze herkend worden? In dit proefschrift waren er doorgaans aanwijzingen zowel voor opslag van gereduceerde vormen in het mentale lexicon (Hoofdstuk 4 en 5) als voor een abstractie over het gereduceerde spraaksignaal die generalisatie-effecten vergemakkelijkt (Hoofdstuk 3 en 4). Gereduceerde woorden lijken tenminste in twee gevallen als uitspraakvarianten te worden opgeslagen. Ten eerste worden ze opgeslagen als hun reductiepatroon zó gevarieerd is dat er geen abstractieregel gegenereerd kan worden (bv. de variabele klinkerdeleties in Hoofdstuk 4). Maar ook gereduceerde woorden die een volkomen regelmatig patroon volgen, kunnen als varianten in het mentale lexicon worden opgeslagen, als luisteraars er zeer vertrouwd mee zijn (bv. de *ge*-reducties in Hoofdstuk 5 waarmee de Zuid-Duitse luisteraars zeer vertrouwd zijn). Wellicht moet deze bewering nog verder worden aangescherpt zodat een regelmatig gereduceerde vorm alleen opgeslagen wordt als a) hij voor de luisteraars ook de natuurlijke uitspraak van het woord is (d.w.z. dat de gereduceerde vorm vóór de canonieke vorm verworven is) of b) luisteraars hem op een natuurlijke manier produceren (d.w.z. de gereduceerde vorm wordt net zo behandeld als de natuurlijke uitspraak, maar hij wordt ná de canonieke vorm verworven). Toekomstig onderzoek zal moeten testen of de leeftijd van verwerving en productie de opslag van regelmatig gereduceerde woordvormen beïnvloedt. Verder lijkt het plausibel dat tenminste sommige regelmatig gereduceerde vormen waarmee de luisteraar niet zeer vertrouwd is (tijdelijk?) in het mentale lexicon opgeslagen worden, zodat er überhaupt abstractieregels kunnen worden gegenereerd. Zou onderzoek daarentegen echter uitwijzen dat leren over reducties gegeneraliseerd kan worden over sprekers, dan zou dit een sterke aanwijzing voor abstractie zijn. Een dergelijke generalisatie is namelijk moeilijk te verklaren met alleen het opslaan van exemplaren van de woorden die door de spreker in de blootstellingsfase zijn geproduceerd.

De resultaten van dit proefschrift, die aanwijzingen leveren voor zowel opslag als voor abstractie, pleiten tegen extreme modellen die slechts één van deze twee mechanismen ten grondslag hebben liggen aan de herkenning van gereduceerde woorden. Van de verschillende hybride modellen kunnen de modellen die geen mechanisme voor generalisatie bevatten (bv., Ranbom & Connine, 2007) de resultaten van Hoofdstuk 3 en 4 niet verklaren. Verder moeten modellen die een generalisatiemechanisme bevatten van representaties zijn voorzien die groter zijn dan een segment zodat ze de adaptatie aan *ver*-reducties die in Hoofdstuk 3 gevonden werd kunnen verklaren. De resultaten van Hoofdstuk 4 suggereren dat deze eenheden, die groter zijn dan een segment, vaststaande sequenties van segmenten moeten zijn, die in vele woorden voorkomen. Het is dus waarschijnlijk dat alleen morfemen aan deze beperking voldoen.

De vele manieren waarop luisteraars zich aan reducties aanpassen

Dit proefschrift liet zien dat luisteraars zich aan gereduceerde woorden kunnen aanpassen, zodat ze dezelfde of andere gereduceerde woorden beter kunnen herkennen. Welke woorden vervolgens het best herkend worden, hangt ervan af hoe uniform gereduceerd de woorden waren die de luisteraars zijn tegengekomen. Als ze volgens een zeer uniform patroon gereduceerd waren, konden luisteraars zeer goed omgaan met soortgelijke vormen van reductie. Als ze volgens een zeer variabel patroon gereduceerd waren, tolereerden luisteraars zeer verschillende vormen van reducties. Ze waren dan flexibel genoeg om zich aan gereduceerde spraak in het algemeen aan te passen. Flexibele perceptuele aanpassingen kunnen zwakkere effecten opleveren dan leren over een uniform reductiepatroon (zie Hoofdstuk 4), maar ze zijn toepasbaar op een grotere bandbreedte van gereduceerde woorden. Luisteraars kunnen dus wel met reducties omgaan, en kunnen dus een fenomeen aan dat zeer frequent is in de dagelijkse spraak.

Curriculum Vitae

Katja Poellmann was born in 1982 in Marktredwitz, Germany. From 2002 to 2004, she studied Mathematics and Computer Science at the University of Bayreuth. She then changed directions and studied Romance languages and literature (French, Italian, German) at the University of Bamberg. Upon completion of her *Vordiplom* in 2007, she was awarded a three-year student scholarship by the renowned *Studienstiftung des deutschen Volkes* (German National Academic Foundation). Katja spent the first six months of 2007 in France as an intern at the *Centre Culturel Franco-Allemand* in Nantes. During the second half of 2007, she studied at the Institute of Linguistics at the University of Pisa, Italy, as an Erasmus exchange student. She received her *Diplom* (equivalent to a Master's degree) in the spring of 2010 and a week later she started as a Ph.D. student in the Language Comprehension Department of the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands. In 2012, she spent four months at the Eberhard Karls University Tübingen (Germany) to work with Prof. R. Harald Baayen. This collaboration was funded by a NetWordS travel grant from the *European Science Foundation*. In January 2014, she will join the Department of Speech-Language Pathology & Audiology at Northeastern University in Boston, USA.

List of publications

Poellmann, K., Mitterer, H., & McQueen, J. M. (submitted). Use what you can: Storage, abstraction processes and perceptual adjustments help listeners recognize reduced forms. *Frontiers in Psychology*.

Poellmann, K., McQueen, J. M., & Mitterer, H. (under revision). Familiarity with regional variation in speaking style is necessary for effortless recognition of morphologically complex words in casual speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Poellmann, K., Bosker, H. R., McQueen, J. M., & Mitterer, H. (under revision). Perceptual adaptation to segmental and syllabic reductions in continuous spoken Dutch. *Journal of Phonetics*.

Poellmann, K., McQueen, J. M., & Mitterer, H. (2011). The time course of perceptual learning. In W.-S. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011 [ICPhS XVII]* (pp. 1618-1621). Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong.

Weber, A., & Poellmann, K. (2010). Identifying foreign speakers with an unfamiliar accent or in an unfamiliar language. In *New Sounds 2010: Sixth International Symposium on the Acquisition of Second Language Speech* (pp. 536-541). Poznan, Poland: Adam Mickiewicz University.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*

20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Danielle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties: ERP studies on the time course in language comprehension. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from nonconstrative variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch: Probabilistic effects in production and perception. *Mark Pluymaekers*

44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: an information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*

67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language: Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: How a village sign language of Bali inscribes its signing space. *Connie de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt J. Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word productions. *Zeshu Shao*
79. Hearing and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representations and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*