Journal of Phonetics 46 (2014) 101–127

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Perceptual adaptation to segmental and syllabic reductions in continuous spoken Dutch

Katja Poellmann [a,b,*], Hans Rutger Bosker [a], James M. McQueen [a,c], Holger Mitterer [a,1]

[a] Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands
[b] International Max Planck Research School for Language Sciences, P.O. Box 310, 6500 AH Nijmegen, The Netherlands
[c] Behavioural Science Institute and Donders Institute for Brain, Cognition & Behaviour, Radboud University Nijmegen, Postbus 9104, 6500 HE Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 8 March 2013
Received in revised form
18 June 2014
Accepted 24 June 2014

Keywords:
Segmental reduction
Syllabic reduction
Adaptation

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/ → [ʊ]) 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. In one of two experiments, the segmental reduction exposure group was better than the syllabic reduction exposure group in recognizing new reduced /b/-words. In both experiments, 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, and hence 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.

1. 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 ([ʊ]) or shorten the prefix *ver-* in *verlangen* "desire" to [fː]. 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/ → [ʊ]) or syllabic (*ver-* → [fː]) reductions, participants might be able to transfer their acquired knowledge about the speaker's pronunciation habits to previously unheard words.

1.1. Compensation for reduction: Multiple mechanisms

In casual speech, a variety of reduced forms can be found ranging from assimilations (e.g., [tœy**m**bɑŋk] for Dutch *tuinbank* "garden bench"; see e.g., Mitterer, Csépe, & Blomert, 2006; Mitterer, Csépe, Honbolygo, & Blomert, 2006) 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;

* Corresponding author. *Present address*: Northeastern University, Speech-Language Pathology & Audiology, 360 Huntington Avenue, 226 FR, Boston, MA 02115, USA. Tel.: +1 617 373 5795.
E-mail addresses: k.poellmann@neu.edu (K. Poellmann), hansrutger.bosker@mpi.nl (H.R. Bosker), j.mcqueen@pwo.ru.nl (J.M. McQueen), holger.mitterer@um.edu.mt (H. Mitterer).
[1] *Present address*: Department of Cognitive Science, Faculty of Media and Knowledge Sciences, University of Malta, Msida MSD2080, Malta.

0095-4470/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2014.06.004

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).

Listeners have different ways of dealing with reduced speech. For example, they can take the (broader) sentence context and the (narrower) phonological context into account. Furthermore, they can extract information about the reduced word itself, such as lexical knowledge, word frequency and frequency of the surface form, from the mental lexicon. Finally, they can make use of fine phonetic detail in the speech signal and even use auditory processes to recognize reduced words.

The sentence context plays an important role in recognizing strongly reduced forms (e.g., Brouwer, Mitterer, & Huettig, 2013; Ernestus, Baayen, & Schreuder, 2002), while the utilization of probabilistic cues from the surrounding context helps listeners to deal with segmental reductions (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.

Moreover, listeners use the knowledge stored in the mental lexicon to recover from reductions. For instance, listeners showed a lexical bias when judging the presence or absence of word-final /t/ in Dutch (Mitterer & Ernestus, 2006). Furthermore, word frequency appears to be one of the major cues that listeners use to compensate for extreme reductions like [ɛik] (e.g., Ernestus, 2000). For example, Mitterer and Russell (2013) showed that listeners exploit the correlation of lexical frequency and phonological reduction and assume, for instance, that a reduced prefix is more likely to belong to a high-frequency word than to a low-frequency word. Connine, Ranbom, and Patterson (2008) emphasize the role of lexically stored pronunciation variants in recognizing schwa-deleted words. In syllable judgment and lexical decision tasks, listeners reacted faster to more frequent surface forms, regardless of whether they were schwa-bearing or not.

Fine phonetic detail is another means by which listeners can recognize reduced words. It can hint at the presence of an apparently deleted segment, like a schwa: Listeners are able, for example, to differentiate a reduced form [spʰɔːt] of "support" from the unreduced form [spɔːt] "sport" (Manuel, 1992). In the case of assimilated forms like [tœy**m**bɑŋk], the intended word can be recognized 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*.

Finally, the fact that assimilated forms can be recognized when listening to an unfamiliar language (Mitterer et al., 2006; Mitterer, Csépe, Honbolygo, et al., 2006) suggests that language-independent perceptual processes are involved in compensation for at least some forms of assimilation. Mitterer et al. (2006) argued further that language-universal auditory processes are involved in compensation, since similar effects could be obtained with speech and non-speech materials.

1.2. Speaker-dependent properties

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 phonetic and language-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 in order to compensate for the speakers' imperfect articulation. 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. Please note that, in the following, we will use the terms *adaptation* and (*perceptual*) *learning* interchangeably.

1.3. Speaker-specific learning

There is ample evidence that listeners tune in to characteristics of particular speakers. For instance, Ladefoged and Broadbent (1957) and Sjerps, Mitterer, and 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