



Research Article

Effects of phonetic reduction and regional dialect on vowel production

Cynthia G. Clopper^{a,*}, Jane F. Mitsch^a, Terrin N. Tamati^b^a Ohio State University, USA^b University Medical Center Groningen, The Netherlands

ARTICLE INFO

Article history:

Received 25 June 2015

Received in revised form

8 September 2016

Accepted 21 November 2016

Available online 16 December 2016

Keywords:

Phonetic reduction

Dialect variation

Lexical neighborhood density

Discourse mention

Speaking style

ABSTRACT

Many linguistic factors contribute to variation in vowel dispersion, including lexical properties, such as word frequency, and discourse properties, such as previous mention. Indexical factors, such as regional dialect, similarly contribute to spectral vowel variation in production. A handful of previous studies have further suggested that linguistic and indexical factors interact such that talkers produce more extreme sociolinguistic variants in linguistic contexts that promote phonetic reduction, such as high frequency and high predictability words. The goal of the current study was to extend the empirical base of this research through an exploration of the interactions between regional dialect and lexical phonological similarity, discourse mention, and speaking style, respectively, on vowel production in Northern and Midland American English. The results revealed more extreme regional dialect variants in reduction-promoting contexts, consistent with previous research. However, substantial variability in phonetic reduction and its interaction with dialect variation was also observed across linguistic contexts, vowel categories, and acoustic domains (temporal vs. spectral), suggesting that a more complex account of the cognitive, linguistic, and indexical factors contributing to phonetic reduction processes is necessary.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Phonetic reduction processes are realized on vowels in both the temporal domain, as shorter duration, and in the spectral domain, as centralization in the vowel space. Although phonetic vowel reduction is considered a distinct process from phonological vowel reduction (i.e., alternations with /ə/; Fourakis, 1991; Padgett & Tabain, 2005), phonetic vowel reduction is typically described as involving centralization towards a “neutral” vowel (i.e., /ə/; Fourakis, 1991; van Bergem, 1993). This centralization is captured quantitatively as the distance in the vowel space from a particular vowel to the acoustic center of the space, defined as the grand mean of F1 and F2 values across vowels (Bradlow, Torretta, & Pisoni, 1996; Fourakis, 1991; Munson & Solomon, 2004; Scarborough, 2010; Scarborough & Zellou, 2013; Wright, 2004; Zellou & Scarborough, 2015). This dispersion approach to assessing spectral vowel reduction reflects the theoretical understanding of this process as involving articulatory undershoot or gestural economy in reduced forms relative to unreduced forms (Lindblom, 1990; Recasens, 1999).

Numerous linguistic contexts have been shown to contribute to phonetic vowel reduction in the temporal and/or spectral domains, including lexical frequency, lexical phonological similarity, semantic predictability, discourse mention, and speaking style (for a review, see Clopper and Turnbull (in press)). Specifically, high frequency words, words with few phonological neighbors, predictable words, repeated mentions of words, and words in a casual speaking style are phonetically reduced relative to low frequency words, words with many phonological neighbors, less predictable words, first mentions of words, and words in a careful speaking style. The theoretical interpretation of these effects is that phonetic reduction arises when the lexical target is easy for the talker and/or the listener to access. For example, Lindblom's (1990) hyper- and hypospeech (H&H) theory describes phonetic reduction as a trade-off between gestural economy, which benefits the talker, and the maintenance of perceptual distinctiveness, which benefits the listener (see also Aylett and Turk (2004, 2006)). That is, according to the listener-oriented account (e.g., Aylett & Turk, 2004, 2006; Lindblom, 1990), phonetic reduction is a result of the competing needs of the talker to maintain comprehensibility of the message while expending as little energy as possible. Thus, when the listener is expected to easily understand the message, the talker can reduce

* Correspondence to: Department of Linguistics, Ohio State University, 1961 Tuttle Park Place, Columbus, OH 43210, USA. Fax: +1 614 292 8833.
E-mail address: clopper.1@osu.edu (C.G. Clopper).

or hypoarticulate, but when the listener is expected to have more difficulty understanding the message, the talker must speak more clearly and avoid reduction. This listener-oriented approach makes strong assumptions about the talker's ability to track the listener's needs. However, the evidence for this type of real-time listener modeling is somewhat mixed (Arnold, 2008; Bard et al., 2000; Kahn & Arnold, 2015; Snedeker & Trueswell, 2003), leading to alternative proposals that phonetic reduction reflects primarily speech production processes, such as retrieval and planning (e.g., Arnold, Kahn, & Pancani, 2012; Baese-Berk & Goldrick, 2009; Bard et al., 2000; Jurafsky, Bell, Gregory, & Raymond, 2001). That is, according to the talker-oriented account (e.g., Baese-Berk & Goldrick, 2009; Jurafsky et al., 2001), phonetic reduction is a result of the cognitive processes underlying speech production. Specifically, word forms that are easy to access in production are reduced relative to word forms that are more difficult to access, due to differences in lexical access speed and/or activation.

In both the listener-oriented and talker-oriented approaches to phonetic reduction, the typical assumption is that the observed effects across all of the various linguistic contexts derive from a single underlying process. For example, Lindblom (1990) drew explicit connections between lexical sources of reduction, such as lexical frequency and phonological similarity, and stylistic sources of reduction. Similarly, Aylett and Turk (2004) drew explicit connections between lexical sources of reduction, including lexical frequency, and discourse-level sources of reduction, including discourse mention. The talker-oriented proposals similarly range across linguistic contexts, including phonological similarity (Baese-Berk & Goldrick, 2009), discourse mention (Bard et al., 2000), and semantic predictability (Arnold et al., 2012; Jurafsky et al., 2001), consistent with the suggestion that the findings across contexts reflect a single underlying process. However, the results of several recent studies provide preliminary evidence that phonetic reduction may arise from different processes (Turnbull, 2015; Watson, 2010). For example, Lam and Watson (2010) found that discourse mention led primarily to temporal reduction, whereas semantic predictability led primarily to reduction in intensity. Watson (2010) interpreted these results as suggesting that intensity and f_0 reflect listener-oriented processes, such as semantic predictability, whereas duration reflects talker-oriented processes, such as priming across discourse mentions. This proposed alignment between acoustic-phonetic domains (duration vs. intensity and f_0), cognitive reduction processes (speech production constraints vs. listener modeling), and linguistic contexts (discourse mention vs. contextual predictability) provides an interesting avenue of inquiry for vowel reduction, especially given that temporal and spectral vowel reduction are not necessarily correlated (Fourakis, 1991; cf. Moon & Lindblom, 1994). In the current study, we consider variability within and across linguistic contexts in temporal and spectral vowel reduction, and their interaction with social indexing of regional dialect information.

1.1. Phonetic vowel reduction and social indexing

The results of a small set of studies from different laboratories have provided preliminary evidence that the same contexts that lead to phonetic reduction also lead to more frequent and/or more extreme sociolinguistic variants that index group identity information about the talker, such as greater /æ/-raising in the Northern dialect of American English or greater /qj/ monophthongization in African American English (Clopper & Pierrehumbert, 2008; Hay, Jannedy, & Mendoza-Denton, 1999; Munson, 2007a). The linguistic contexts examined in these studies range from lexical properties to discourse properties and the relevant social categories include regional background, racial identity, and sex typicality. In the first of these studies, Hay et al. (1999) examined the production of /qj/ in the speech of Oprah Winfrey, an African American talk show host. They observed more monophthongal /qj/ tokens, consistent with Ms. Winfrey's African American identity, in high frequency words relative to low frequency words. Similarly, Munson (2007a) observed a larger sex difference in vowel space dispersion for high frequency and low neighborhood density words than for low frequency and high neighborhood density words, respectively. Finally, Clopper and Pierrehumbert (2008) observed more extreme regional dialect variants for Northern American English talkers in words that were predictable given their preceding sentence context relative to words that were less predictable. Thus, across all three of these reduction-promoting contexts (high frequency, low density, and high predictability targets), more extreme sociolinguistic variants and/or larger sociolinguistic differences between groups were produced relative to contexts that do not promote phonetic reduction.

The theoretical account of these findings follows Lindblom's (1990) listener-oriented approach. In particular, all three sets of authors argued that talkers produce more extreme sociolinguistic variants in contexts in which they are likely to be understood by the listener (Clopper & Pierrehumbert, 2008; Hay et al., 1999; Munson, 2007a). That is, just as easy contexts for the listener allow for gestural economy leading to reduction, easy contexts for the listener allow for more extreme social indexing without sacrificing communication. Critically, this listener-oriented interpretation of the interaction between phonetic reduction and social indexing does not assume that social indexing is a conscious process. Hay et al. (1999) made this point explicitly in their discussion of Ms. Winfrey's speech, but Munson's (2007a) data also provide evidence for the implicit nature of this interaction. In particular, Munson's (2007a) results show a larger sex-based difference in vowel space dispersion for low density words than for high density words. This interaction is unlikely to be the result of an explicit or conscious strategy on the part of the talkers in his study to exaggerate sex differences in low density words, given that vowel space dispersion is not stereotypically associated with sex differences and that native speakers do not have strong intuitions about lexical neighborhood density. Further, although the listener-oriented perspective is directly invoked in all three of these previous studies, talker-oriented factors related to processing time and the nature of socially-indexed representations may also contribute to the observed findings. For example, more extreme sociolinguistic variants may be easier to access for the talker, leading to the production of those variants in easy processing contexts (Clopper & Pierrehumbert, 2008; Munson, 2007a). Thus, in parallel to the phonetic reduction literature, both listener-oriented and talker-oriented accounts of the interaction between phonetic reduction and social indexing have been proposed.

Download English Version:

<https://daneshyari.com/en/article/5124106>

Download Persian Version:

<https://daneshyari.com/article/5124106>

[Daneshyari.com](https://daneshyari.com)