Journal of Phonetics 39 (2011) 312–318

Contents lists available at ScienceDirect

Journal of Phonetics

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

The effect of reduction on the processing of flaps and /g/ in isolated words

Benjamin V. Tucker *

Department of Linguistics, University of Alberta, 4-32 Assiniboia Hall, Edmonton, Alta., Canada T6G 2E7

ARTICLE INFO

Article history:
Received 6 February 2009
Received in revised form
2 December 2010
Accepted 2 December 2010
Available online 15 January 2011

ABSTRACT

This study investigates the comprehension of words with reduced and unreduced flaps and /g/s by means of an auditory lexical decision task. The experimental stimuli are compared to words produced in various speech styles confirming their naturalness and illustrating the distribution of the experimental stimuli. The experiment shows that reduced words are more difficult to process than unreduced words, and that listeners' response latencies for unreduced words are longest for the highest and lowest frequencies. Listeners' error rates are highest for the lowest frequency items. A comparison of the flap stimuli to the /g/ stimuli reveals no difference in processing between the two phonemes. When an acoustic measure (change in intensity) is used as a predictor, instead of a dichotomous factor contrasting reduced with unreduced, items on the extreme ends of the acoustic measure (very reduced or very unreduced) have longer response latencies than the middle items. The results of the second analysis supports the view that listeners optimize comprehension for intermediate degrees of reduction. The data also suggests that production frequency, or predictability (e.g., the likelihood of a reduced flap) affects the comprehension of reduced and unreduced words, such that highly predictable productions have shorter response latencies.

1. Introduction

Spontaneous speech used in day-to-day interactions is a highly complex and variable phenomenon. A growing body of research is investigating natural, conversational and casual types of speech, the types of speech encountered in normal, daily communicative exchanges (e.g., Arai, 1999; Ernestus et al., 2002; Johnson, 2004; Jurafsky et al., 2001; Mehta and Cutler, 1988; Pickett and Pollack, 1963; Pluymaekers et al., 2005, see also this volume, among others). One focus of these studies has been to investigate the processing of phonetic reduction found in spontaneous speech. For example, some of these studies have investigated word-final /t/ variation in English (e.g., Guy, 1980; Neu, 1980; Sumner and Samuel, 2005) and Dutch (e.g., Janse et al., 2007; Mitterer and Ernestus, 2006), where the segment may be deleted or glottalized in a particular context. Words containing a reduced or deleted segment, however frequent they may be in speech, are more difficult to process.

Word frequency and predictability have been shown to play a major role in modeling reduction of spontaneous speech production (Aylett and Turk, 2004, 2006; Jurafsky et al., 2001; Pluymaekers et al., 2005). For example, if words are more predictable based on their frequency of occurrence or their probability given the following word, they tend to be more reduced. However, in the comprehension of reduced speech, word frequency appears to play a much smaller role. Ernestus and Baayen (2007) find, in an investigation using duration as a measure of reduction, that frequency has a strong effect on the processing of unreduced words. However, they find in two experiments a weak or no effect of frequency for reduced words. They argue that productions of high frequency words are more likely to show reduction due to well learned articulatory motor programs. They refer to this as a speaker-driven explanation (see also Bybee, 2001). Their data challenge listener-driven approaches, as proposed by Aylett and Turk (2004, 2006), in which speakers adjust their speech to facilitate processing for the listener by providing a smooth signal. One might expect that the signal would remain smooth for the listener due to frequency compensating for reduction. In other words, a frequency effect for reduced words is expected, contrary to the findings of Ernestus and Baayen (2007).

The present study examines word frequency as a predictor in comprehension to further compare listener- and speaker-driven approaches. More specifically, the comprehension of reduction found in word-medial flaps and /g/ is investigated using an auditory lexical decision task.

Flaps have been the focus of many previous studies which address the processing and distribution of the careful [t^h] and canonical [ɾ] pronunciation variants (e.g., Charles-Luce, 1997; Connine, 2004; McLennan et al., 2003). For example, *atom* is usually pronounced as [æt^həm] in North America instead of the careful pronunciation [æt^həm], like one may find in dictionaries as the standard British pronunciation. Patterson and Connine (2001) observed that in two corpora flaps are the most frequent production variant, occurring over 93% of the time. McLennan et al. (2003) in six repetition priming studies argued that their priming results for both flaps and careful pronunciations indicated that the flap was stored in the lexicon.

* Tel.: +1 780 492 5952; fax: +1 780 492 0806.
E-mail address: bvtucker@ualberta.ca

0095-4470/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2010.12.001

Contrasting the careful [t] to the flap, Connine (2004) found that listeners were more likely to identify phonemes that formed a word (e.g., identify /p/ for *pretty* as opposed to *bretty*) in a flap environment than in an environment with a careful pronunciation. Connine claimed that these results indicate that the flap is part of the lexical representation. These results taken together provide evidence that the flap is indeed stored in the lexicon. Ranbom and Connine (2007) find that the processing careful /nt/ productions elicit the fastest response times and that the more frequent production variant, the nasal flap [ɾ̃] (as found in words like *gentle*), elicited a gradient response pattern, such that the higher production frequency items, elicited faster and more accurate responses than the less predictable nasal flap items. They claim that this result provides evidence that the mental lexicon contains at least an orthographic representation and a gradient representation of nasal flaps based on production frequency (the likelihood of an allophone variant being produced).

Unlike the previous research on alveolar flaps, this paper investigates flap reduction, a process of flaps being produced like an approximant. For example, *puddle* is pronounced with a word-medial canonical or unreduced flap as [pʌɾl] but is also commonly produced in spontaneous speech as [pʌɾ̞l], with an approximant-like variant of the flap (illustrated in Fig. 1). Word-medial /g/ items were selected to act as a comparison group to the flap items (/d/ phonemes) because they do not undergo flapping. However, like flap reduction, word-medial /g/ can be reduced to an approximant-like variant (e.g., [bægi] becomes [bæg̞i]). A comparison of the reduced forms of /d/ and /g/ is potentially informative about the status of the flap in the lexicon. If the flap is stored in the lexicon (as predicted by the research described above), then both the flap and the /g/ undergo the same reduction process and no difference is predicted in the response latencies. However, if the reduced form is derived from an underlying /d/, then two processes are involved and longer response latencies are expected as compared to the /g/.

A first experiment investigating the comprehension of words with reduced /g/ and flap made use of a cross-modal identity priming paradigm Tucker and Warner (2007). Tucker and Warner (2007) used a subset of the materials in the present study. The results of Tucker and Warner (2007) indicate that processing of reduced items was inhibited, stronger priming for unreduced items (shorter response latencies) than for reduced items. They also found that flap items were responded to faster (significant for subjects but not for items) and more accurately (significant for subjects and items) than /g/ items, providing marginal support that the flap is stored in the lexicon. One disadvantage of the cross-modal identity priming task is that participants respond to orthographic targets. If the orthographic form of the word is stored in the lexicon (e.g., Ranbom and Connine, 2007), then listeners may have been biased by the target toward a more canonical pronunciation (the unreduced form) of the flap or /g/. An auditory lexical decision task avoids any potential experimentally induced orthographic bias.

The present study complements Tucker and Warner (2007) with a detailed description of the materials used and extends the findings by investigating possible interactions of word frequency and word duration in the comprehension of reduced word forms. In sum, the following questions are investigated in this paper: (1) Does lemma frequency play a role in the comprehension of unreduced segments but not for the reduced segments Ernestus and Baayen (2007)? (2) Is comprehension affected differently by reduction in the flap than in the /g/? (3) Does reduction inhibit comprehension of word-medial flap and /g/ or is this simply the result of an orthographic bias?

2. Experiment

2.1. Participants

Participants were 64 native speakers of American English, recruited at the University of Arizona. Listeners consisted mainly of undergraduate students enrolled in Introduction to Linguistics courses. All received extra credit for participating. None reported any known hearing loss. Listeners' ages ranged from 18 to 60 years. Over 50% of the listeners were either studying a foreign language (generally Spanish or French), or were late bilinguals. Nineteen of the listeners were self-reported monolinguals.

2.2. Materials

Forty experimental items contained flaps (orthographically 'd') and forty contained word-medial /g/ (orthographically 'g'). Only orthographic 'd' items were used for the flaps to reduce any potential competing effects of orthography within the experiment from 't' flapped items. All items were bisyllabic and had first syllable stress.

In addition to the 80 real-word experimental items there were 220 fillers (200 non-word fillers and 20 real-word fillers) and 10 additional practice items. Twice as many non-words were used so that no real-word was preceded by another real-word, in order to prevent any 'Yes' biased responses. All non-words were bisyllabic and phonotactically possible English words.

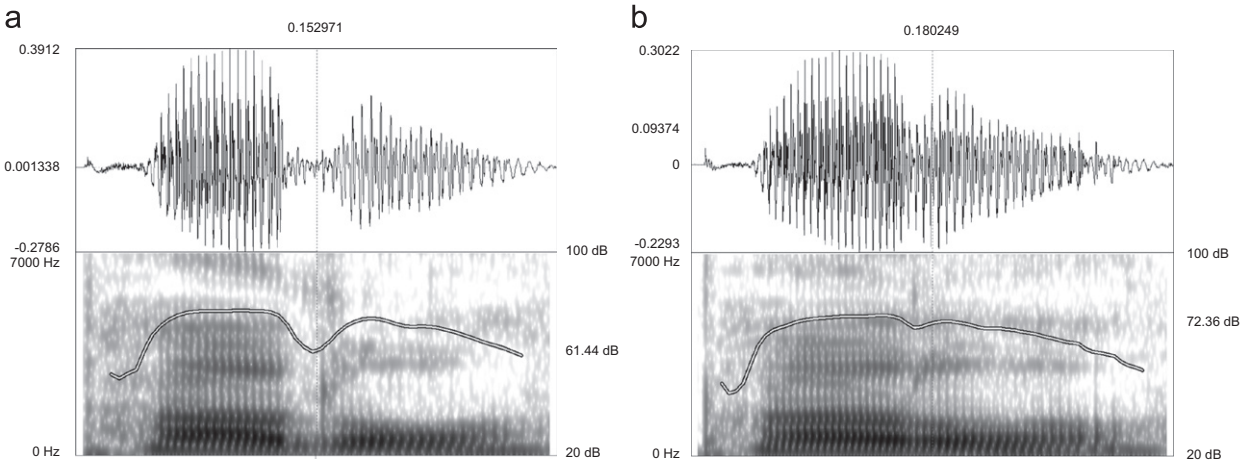

Fig. 1. Waveform and spectrogram of the unreduced (a) and reduced (b) productions of the item *puddle*. The line represents the intensity of the sound overlaid on the spectrogram.