Journal of Phonetics 71 (2018) 194–214

Contents lists available at ScienceDirect

Journal of Phonetics

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

Research Article

Relative cue weighting in production and perception of an ongoing sound change in Southern Yi

Jianjing Kuang *, Aletheia Cui

Department of Linguistics, Suite 300, C-3401 Walnut Street, Philadelphia, PA 19104, United States

ARTICLE INFO

Article history:
Received 11 October 2017
Received in revised form 7 September 2018
Accepted 8 September 2018

Keywords:
Sound change
Cue weighting
Production and perception

ABSTRACT

Multiple co-varying cues for a phonological contrast are often introduced by coarticulation, and sound change occurs when their relative weighting shifts. The central issues for this kind of sound change include how cue weighting shifts over time in both production and perception and what the mapping is between production and perception during this process. This study aims to provide insights to these questions by examining an ongoing change in the tense vs. lax register contrast in Southern Yi. Production and perception experiments were conducted with the same group of speakers to evaluate the relative importance of the source cue (i.e., phonation) and its coarticulated cues (i.e., vowel formants and f0) for this contrast. While speakers of all age groups still maintain the register contrast, our results show that formant differences are overtaking phonation as the primary cues. This sound change is more advanced in non-high vowels than high vowels in both perception and production. Moreover, production and perception are misaligned; in both cases, the shift to formant values occurs first in perception, with production lagging behind. These findings illustrate the nuanced progression of sound change and a better understanding of the role of production and perception in the initiation of a sound change.

1. Introduction

Phonological contrasts are usually realized with multiple co-varying cues, and sound change can occur when the primary cue of the given contrast shifts. Tonogenesis is a well-known case for this type of sound change – while the original lexical contrasts are maintained, pitch overtakes other cues such as voicing as the primary cue. An important theoretical question about this type of sound change is how and when the primary cue shifts in production and perception. In particular, do speakers or listeners lead the shift? What is the interaction between production and perception during the cue shifting in progress? The goal of this study is to provide insights to these questions by examining an ongoing cue-shifting change in Southern Yi, where vowel quality is overtaking phonation as the primary cue for its register contrast. This understudied sound change is analogous to tonogenesis in many ways, and thus can provide us with an important case study to better understand the initiation stage of this type of sound change *in vivo*.

1.1. Cue shifting and sound change

1.1.1. The role of co-varying cues in sound change

Speech signals are highly redundant. In speech production, an articulatory target is often achieved by the coordination of multiple articulators. As a result, this process, known as coarticulation,[1] is one of the major causes for a given phonological contrast to have multiple co-varying acoustic cues. For example, Lisker (1986) noted that the voiced vs. voiceless contrast of English obstruents involves at least 16 acoustic cues, including the intensity of the glottal signal, the duration of the vowel, the duration of the first formant transition, F1 offset frequency, voice onset time (VOT), f0 contour, and so on. While listeners take advantage of multiple cues to ensure the success of perceiving the intended linguistic targets (e.g., Brunelle, 2012; Kingston, Diehl, Kirk, & Castleman, 2008; Kuang, 2013; Toscano & McMurray, 2010), the co-varying cues differ in their contribution, or weights, to a phonological contrast, as listeners' attention is selective. For example, among the large set of co-varying

* Corresponding author.
E-mail address: kuangj@ling.upenn.edu (J. Kuang).

[1] Coarticulation is also commonly used to refer to the contextual influence from neighboring segments.

https://doi.org/10.1016/j.wocn.2018.09.002
0095-4470/

cues, voice onset time (VOT) has been found to be the most reliable acoustic cue for the English voicing contrast in the onset position (Lisker & Abramson, 1964, 1970; Schertz, Cho, Lotto, & Warner, 2015; Francis, Kaganovich, & Driscoll-Huber, 2008; Davidson, 2016; Kong & Edwards, 2016; Nelson & Wedel, 2017).

Nonetheless, secondary cues can provide opportunities for sound change, which may occur when the relative weighting of the cues shifts. One of the best-known instances of this kind of sound change is tonogenesis. It has been well-established that vowels following a voiceless onset tend to have higher f0 (e.g., Hombert, Ohala, & Ewan, 1979; Kingston, 2005; Maddieson, 1984; Ohala, 1973; Thurgood, 2002), and this f0 perturbation effect is fairly common among languages (see Hombert et al., 1979 for Yoruba, Xu & Xu, 2003 for Mandarin, Haggard, Ambler, & Callow, 1970 for English, and Jun, 1996 for Korean and French). Notably, in some languages, such as Vietnamese (Thurgood, 2002), this synchronic variation in f0 eventually developed into phonemic contrasts, i.e., tones. In this kind of change, while a phonological contrast is preserved, one cue overtakes another as primary. Before tonogenesis, the primary contrastive cue is the voicing of the onset consonant. After tonogenesis, the primary cue for the contrast has shifted to the pitch differences on the vowels, and the voicing contrast on the consonants is often completely lost (Hyman, 1976; Hombert et al., 1979; Kingston, 2005; Kirby, 2013; Thurgood, 2002). An important question is how a coarticulated cue rises in significance and becomes the primary cue.

According to Hyman (1976), tonogenesis involves three steps: (1) voiced and voiceless consonants determine the f0 perturbations on following vowel as the result of intrinsic coarticulation, (2) f0 perturbations are exaggerated and become a perceptual cue (i.e., phonologized), and (3) distinct tones develop and the consonant voicing distinction is lost (i.e., phonemicization). In this proposal, phonologization in step 2 is crucial for synchronic variation to turn into sound change, during which an acoustic cue (e.g., f0) becomes a significant but secondary contributor in both the production and the perception of voicing. However, this step by itself is not enough to trigger sound change (e.g., tonogenesis), because it is natural for languages to have multiple stable secondary cues (Wright, 2004) as enhancement cues for the target contrast (Stevens & Keyser, 2010). For example, f0 (Haggard, Kleber, & Reubold, 1970), F1 formant transition (Liberman, Delattre, & Cooper, 1958), vowel duration (Summerfield, 1981), aspiration amplitude (Repp, 1979), and burst spectrum (Chodroff & Wilson, 2014) all significantly contribute to the voicing contrast in both production and perception for English speakers, but English is not undergoing tonogenesis. Actual tonogenesis occurs only when the contrast is reanalyzed and the primary cue of the contrast shifts from VOT to f0. In other words, there must be additional intermediate steps between step 2 and step 3, where the novel phonologized secondary cue rises in significance and becomes the primary cue in production and perception. The question is, then, how is the change of the primary cue implemented in production and perception?

1.1.2. The mapping between perception and production during sound change

In theory, there are three possibilities for the time course of cue shifting:

(1) the primary cue shifts in production and perception at the same time,
(2) listeners first shift their attention to a new cue in perception, and then in turn rely on this cue to mark a phonological contrast in production, and
(3) cue shifting starts in production, and listeners subsequently become attuned to the changes in perception.

Possibility (1) assumes parity between production and perception during sound change. Generally speaking, there should be parity between production and perception because gestures can usually be recovered from the speech signal (Fowler, 2005; Fowler & Smith, 1986). Both possibility (2) and possibility (3) assume that there can be misalignment between production and perception, i.e., speakers rely on different primary cues in production and perception.

Several models of sound change support the scenario in possibility (2). Ohala's model of sound change (Ohala, 1981, 1993) and the extensions thereof (e.g., Solé, 2014) proposed that the driving force of sound change is the unintentional error on the part of the listener. In these proposals, sound change occurs when the listener fails to compensate for the effects of contextual coarticulation or when they attribute the coarticulated effects to a wrong source. While also recognizing that listeners are the driving force of sound change, Beddor (2009) suggested that this kind of parsing is not a mistake on the listener's part. Rather, listeners actively attend to any relevant cues because multiple grammars are consistent with the input. Some listeners simply place more weight on the effect than on the source of the coarticulation. Beddor (2012) also hypothesized that variation in interpreting the cues is earlier than the actual sound change, which is a process that results in the emergence of the innovative cues. Altogether, these theories predict that the reanalysis of the contrastive cues (i.e., the shift of the primary cue) is manifested in perception before production. It should be noted that, these theories also assume that before perceptual reanalysis, the raw materials for cue shifting, such as reliable secondary cues, are already grounded in production.

While possibility (2) states that perception plays an active role in reanalysis, possibility (3) suggests that perception plays a rather passive role, and that cue shifting is largely driven by changes in production. In particular, it has been proposed (e.g., Abramson, 2004; Kang, 2014; Kirby, 2013) that a secondary cue is likely to take over as the primary cue when the original primary cue is neutralizing and merging – that is, to "save" the lexical contrasts, speakers may emphasize a secondary cue to compensate for the loss of the primary cue. Under this proposal, the shifting of the primary cue might happen in production first. Another reasoning (e.g., Janson, 1983) is based on the fact that speakers who utilize different cue weights in production (e.g., old vs. young speakers) can usually perfectly understand each other. In order to maintain the mutual intelligibility in communication, cue shifting in perception might be slower than in production.