Journal of Phonetics 40 (2012) 162–176

Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

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

Coarticulation between tone and glottal consonants in Itunyoso Trique

Christian T. DiCanio *

Laboratoire Dynamique du Langage, Université de Lyon-2, France

ARTICLE INFO

Article history:
Received 13 May 2010
Received in revised form
21 September 2011
Accepted 5 October 2011
Available online 10 November 2011

ABSTRACT

This paper investigates the realization of contrastive tone in three non-modal phonation contexts (creaky phonation, glottal closure, and breathy phonation) in Itunyoso Trique, an Oto-Manguean language spoken in Oaxaca, Mexico. The study examines how coarticulatory glottalization (creaky phonation, glottal closure) coincides with coarticulatory pitch perturbations and spectral tilt changes on neighboring vowels. The onset of laryngeally induced F_0 perturbation effects and the timing of changes in spectral tilt were examined using acoustic data from six speakers of the language. The results show that in contexts where substantial non-modal phonation spreads onto the adjacent vowel, greater pitch effects are observed. In contexts where abrupt glottal closure occurs, less coarticulatory changes in spectral tilt and pitch are observed on adjacent vowels. In addition, strong tonal effects are observed for certain spectral measures. These findings are discussed in relation to the literature on tonogenesis and coarticulatory pitch effects.

1. Introduction

One of the principal topics of study in phonetics is how sounds are coarticulated with one another in the process of speech production. Coarticulation varies considerably cross-linguistically (Beddor, Harnsberger, & Lindemann, 2002; Farnetani & Recasens, 1999; Gibbon, Hardcastle, & Nicolaidis, 1993; Han, 2007; Maddieson & Emmorey, 1985; Manuel, 1990, 1999; Öhman, 1967; Przezdziecki, 2005; Scarborough, 2004; Whalen, 1990). A substantial amount of research has argued that such variation is not merely language-specific, but the result of more general mechanistic (Browman & Goldstein, 1986, 1990; Byrd, 1996a, 1996b; Gick, 1999; Gick, Campbell, Oh, & Tamburri-Watt, 2006; Krakow, 1989, 1993, 1999) or perceptual constraints on speech production (Chitoran, Goldstein, & Byrd, 2002; Gick et al., 2006; Keyser & Stevens, 2006; Kochetov, 2002, 2006; Silverman, 1997a, 1997b; Silverman & Jun, 1994; Wright, 1996). While such explanations reduce coarticulation to more general linguistic principles, a number of recent cross-linguistic studies argue that coarticulation is not sensitive to perceptual or mechanical factors, but is largely language-specific. Language-specific factors which influence coarticulatory timing include lexical contrast (Scarborough, 2004), phonological inventory size (Manuel, 1999), and the phonological distribution of contrasts (Bird, Caldecott, Campbell, Gick, & Shaw, 2008; Howe & Pulleyblank, 2001; Manuel, 1999; Miller, 2007).

A majority of the research on coarticulation has focused on consonant and vowel timing. Work on the coarticulation of tone[1] or phonation-type contrasts has received relatively less attention. These studies have concentrated mainly on tone–tone coarticulation (Brunelle, 2003; Peng, 1997; Xu, 1994, 1997, 1999, 2004, 2005; Xu & Liu, 2006; Xu & Sun, 2002) and, to a lesser extent, tone–syllable alignment (Maddieson, 2004; Myers, 2000; Xu, 2004, 1998; Xu & Liu, 2006; Zsiga & Nitisaroj, 2007). Yet, few studies address how tones are coarticulated with particular sound types (see Silverman, 1997a, 1997b; Xu & Xu, 2003; Zee, 1980). The present work focuses on how tone is coarticulated with glottal consonants in Itunyoso Trique, an Oto-Manguean language with nine contrastive tones (DiCanio, 2008). The language is ideal for the investigation of tone–glottal coarticulation, because tones are distributed quite freely in contexts with glottal consonants, e.g. on vowels in /Vʔ/, /Vh/, and /VʔVh/ contexts. This distribution allows one to examine the realization of tone in a variety of different contexts. This study serves two main purposes: to expand the empirical basis for the analysis of tone–segment coarticulation and to test how changes in the relative timing of glottal consonants affect F_0 on adjacent vowels.

The study also tests the hypothesis that gradual phasing between vowels and glottal consonants will result in greater coarticulatory effects on F_0 (Silverman, 1997b). Abrupt phasing occurs where the glottal consonants surface with long-duration closure (for a glottal stop), or long-duration voiceless aspiration (for a glottal fricative). Gradual phasing of glottal consonants occurs when they are accompanied with creaky and breathy phonation. Assuming that abrupt transitions are associated with

* Tel.: +33 4 69 60 07 01.
E-mail addresses: cdicanio@gmail.com, dicanio@haskins.yale.edu

[1] In this paper, all references to *tone* refer to lexical pitch contrasts, not intonational units.

0095-4470/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2011.10.006

smaller coarticulatory effects (and less coarticulatory blending) (Farnetani & Recasens, 1999; Munhall & Löfqvist, 1992), one predicts less F_0 perturbation in contexts where glottal consonants are abruptly phased and greater F_0 perturbation effects in contexts where they are gradually phased. Coda glottalization is abruptly phased in Itunyoso Trique, while coda breathiness and intervocalic glottalization are gradually phased.[2]

The article is organized as follows: in Section 1.1, I provide a background on the interaction between glottal consonants, non-modal phonation type, and F_0. In Section 1.2, I provide a short discussion of the predictions in the current study. In Section 1.3, I provide the background of the Itunyoso Trique language. In Section 2, I discuss the methods used for the elicitation and analysis of the spectral tilt and F_0 data. In Sections 3.2–3.4, I present the results for each of the three comparisons. In Section 4, I discuss the results from the three comparisons. The findings show significant F_0 effects with gradual phasing patterns but a lack of F_0 perturbation with abrupt phasing patterns, which confirms the hypothesis above. The timing of coarticulatory non-modal phonation correlates closely with the observed patterns of F_0 perturbation.

1.1. The relationship between non-modal phonation type and F_0

Prior to the investigation of tone–glottal coarticulation, I provide some background on the interaction between glottal states and F_0. Non-modal phonation type and glottal consonants are known for influencing F_0, both synchronically (DiCanio, 2009; Garellek & Keating, 2011; Lee, 2008; Mazaudon & Michaud, 2008; Silverman, 1997b; Watkins, 2002), and diachronically via processes of tonogenesis and tonal mutation (Dürr, 1987; Haudricourt, 1954; Hombert, 1979; Kingston, 2005; Mazaudon & Michaud, 2008; Svantesson & House, 2006; Thurgood, 2002). Glottal closure and glottal frication have been associated with F_0-lowering and F_0-raising on adjacent vowels. Non-modal phonation types, such as breathy, creaky, and tense phonation, also vary in the magnitude to which they perturb F_0. Yet, there is a greater convergence in the literature concerning the relationship between phonation type and F_0 than there is for glottal consonants and F_0.

Coarticulatory F_0 effects associated with glottal closure or frication vary depending on whether the glottal gesture is bound with an oral consonant gesture. For instance, ejective consonants which surface with complete glottal closure (so called *stiff ejectives*) often induce F_0-raising on the following vowel, whereas slack ejectives induce F_0-lowering and creaky phonation on the following vowel (Kingston, 1985). This distinction between ejective types is theoretically appealing, as it categorizes abruptly timed glottal gestures with certain predictable phonetic effects and more gradually timed glottal gestures with different effects. However, this distinction is also problematic. For instance, in Witsuwit'en, ejectives induce F_0-lowering for some speakers (mostly women) and F_0-raising for others (mostly men) (Wright, Hargus, & Davis, 2002). The authors observed F_0-lowering and F_0-raising effects for both longer, prototypically stiff ejectives and for shorter slack ejectives. F_0 lowering and raising patterns are also observed for aspirated obstruents. Higher F_0 is observed after aspirated stops in Korean (Cho, Jun, & Ladefoged, 2002; Dart, 1987; Hardcastle, 1973; Jun, 1993, 1995), Cantonese (Francis, Ciocca, Ka Man Wong, & Ka Lam Chan, 2006), and English (Hombert, 1975; Kingston, Diehl, Kirk, & Castleman, 2008). However, aspirated stops condition F_0-lowering in Mandarin (Xu & Xu, 2003). These findings suggest that one cannot easily predict how ejectives and aspirated obstruents will influence F_0.

For glottal closure gestures unassociated with oral consonant gestures, similar mixed findings are observed. Historically, coda glottal closure induced pitch-raising in Vietnamese tonogenesis (Diffloth, 1989; Haudricourt, 1954; Thurgood, 2002), Chinese tonogenesis (Matisoff, 1973), Molinos and San Miguel el Grande Mixtec tonogenesis (Dürr, 1987), and Kaska and Tanacross Athabaskan tonogenesis (Kingston, 2005). In Burmese, only low tones surface before a coda [ʔ], reflecting a historical pattern of pitch-lowering before coda glottal closure (Lee, 2008). However, glottal closure also induced pitch-*lowering* in Sarcee and Navajo tonogenesis (Kingston, 2005) and in Silacoyoapan and Alacatlazala Mixtec tonogenesis (Dürr, 1987). In Mixtepec Mixtec, historical */ʔ/ induced pitch lowering intervocalically, but pitch raising word-finally (Dürr, 1987). Thus, historically, /ʔ/ may induce pitch lowering or pitch raising effects even within variants of the same language family, e.g. Athabaskan and Mixtecan. Synchronic findings are no less ambiguous. Coda [ʔ] induces F_0 lowering on the preceding vowel in Lhasa Tibetan (Hu & Xiong, 2010) but F_0-raising on the preceding vowel in Arabic (Hombert, Ohala, & Ewan, 1979). In some cases, glottal closure may simply fail to induce *any* consistent F_0 perturbations on adjacent vowels, as in Navajo (DeJong & McDonough, 1993).

The patterns observed for glottal fricatives are different. When it is unassociated with an oral consonant gesture, the typical pattern is for F_0-lowering to occur (Hombert et al., 1979). For instance, it is well-established that Vietnamese falling tones derived from historical coda /h/ (Thurgood, 2002). In Kickapoo (Algonquian), vowels preceding a preaspirated stop or /h/ are produced with a low F_0 (Gathercole, 1983). In Arabic, the vowel preceding a coda /h/ also undergoes F_0 lowering (Hombert, 1976).

The variability in glottally induced F_0 perturbation effects can be partially explained if one considers how glottal gestures can overlap within the speech signal. In contexts where glottal frication does not occur with devoicing, a breathy or lax phonation type may occur. In contexts adjacent to glottal closure or when complete glottal closure is not achieved, creaky or tense phonation may occur. Both breathy and creaky phonation type induce F_0-lowering (Gordon & Ladefoged, 2001). Examples where breathy or lax phonation induce F_0-lowering are found in a variety of languages, such as Chong (DiCanio, 2009; Thongkum, 1991), Wa (Watkins, 2002), Hani (Ladefoged & Maddieson, 1985), Yi (Ladefoged & Maddieson, 1985), Jalapa Mazatec (Garellek & Keating, 2011), and Santa Ana del Valle Zapotec (Esposito, 2004). Absent are cases where breathy phonation causes F_0-raising to occur. Creaky phonation is also associated with significant F_0-lowering in a variety of languages, such as San Lucas Quiaviní Zapotec (Chávez Peón, 2010), Santa Ana del Valle Zapotec (Esposito, 2004), Yucatec Maya (Frazier, 2009), Coatzospan Mixtec (Gerfen & Baker, 2005), Mandarin Chinese (Silverman, 1997b), and within English (Hillenbrand & Houde, 1996; Huffman, 2005). Tense phonation, on the other hand, co-occurs with higher F_0 values (DiCanio, 2009; Ladefoged & Maddieson, 1985).

1.2. Predictions concerning coarticulatory F_0 effects

The greater consistency of F_0 effects associated with phonation type allows us to make certain predictions. If glottal consonants are produced in a gradual manner, with substantial accompanying changes in non-modal phonation type, one predicts to find significant F_0 effects on an adjacent vowel. However, the predictions for abruptly timed glottal consonants are unclear. One possibility, comprehensively discussed in Silverman (1997b), is that glottal consonants are abruptly phased relative to vowels so that they may be perceptually recovered by the listener.

[2] Throughout this paper, *glottalization* is used as a general term for either sustained glottal closure (a glottal stop) or vocal fry. Creaky phonation is used solely to refer to vocal fry (Gerratt & Kreiman, 2001).

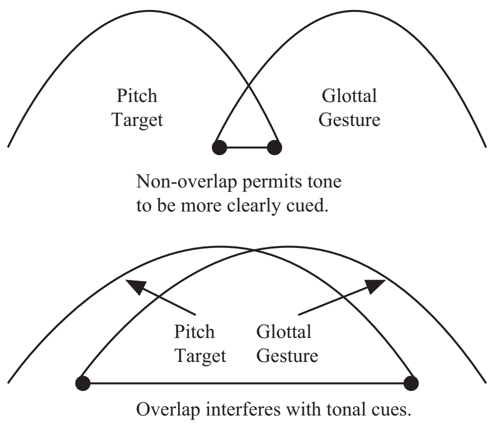

Fig. 1. Coarticulatory overlap governed by perceptual constraints.

In languages with contrastive use of phonation type and tone, or glottal consonants and tone, the tonal and the glottal gestures are abruptly phased so that listeners can distinguish the tone and the glottal contrast. Since phonation type and glottal consonants usually affect F_0, their sequencing preserves optimal tonal contrasts in the language. Explicit in this analysis is the view that tone is optimally perceived on modal vowels. If this is true, then it follows that tone would be timed in many languages so as to not overlap with non-modal phonation type or glottal contrasts. This timing pattern avoids the potential that listeners will misperceive F_0. This idea is schematized in Fig. 1.

This explanation for the timing of F_0 relative to phonation-type or glottal contrasts is compelling, as it accounts for the *phonological* distribution of such contrasts in a large number of languages of the world. While speakers may be generally sensitive to how they may be perceived by listeners, there is less clear phonetic evidence that such knowledge plays a role in determining coarticulatory overlap. In terms of coarticulation, Silverman's theory predicts that abrupt glottal phasing prevents coarticulatory F_0 effects. If glottal consonants are produced in an abrupt manner, without substantial coarticulatory changes in non-modal phonation type, one predicts less F_0 perturbation. The current study directly examines the hypothesis that the magnitude of coarticulatory overlap between glottal consonants and vowels is correlated with the amount of corresponding F_0 perturbation. Apart from testing this relationship, this paper addresses a separate empirical question: how are tones and laryngeals produced in Itunyoso Trique.

1.3. Itunyoso Trique background and phonology

Itunyoso Trique is spoken in the town of San Martín Itunyoso, located in Oaxaca, Mexico. Two other Trique languages are spoken in geographic proximity to Itunyoso Trique: Chicahuaxtla Trique (Good, 1979; Longacre, 1952, 1959) and Copala Trique (Hollenbach, 1977, 1984a, 1984b, 2004, 2007). Like Itunyoso Trique, they too have large phonemic tone inventories where tones co-occur with syllable-final glottal consonants. Background information on Itunyoso Trique is summarized here from DiCanio (2008, 2010), to which the reader is referred for more information.

Unlike most tonal East Asian and Southeast Asian languages, which are monosyllabic in lexical structure, a majority of the lexicon in Itunyoso Trique is composed of polysyllabic words. All such words have final prominence. Prominence is reflected in the phonology via a largely asymmetrical distribution of contrasts which are restricted to the prominent syllable but neutralized in non-final syllables. Prominence is also phonetically realized by increased duration in final rimes. Even though the language contrasts nine tonal patterns, tone is asymmetrically distributed, where a much smaller subset of the contrastive tones in the language may occur on a non-final syllable than on a final syllable. Certain tones have a restricted distribution with respect to the coda glottal consonants /h/ and /ʔ/. Apart from these two glottal consonants, the language lacks codas. Examples of the tonal patterns in the different glottal contexts are given in Table 1.

Table 1
Tones on monosyllables in laryngeal contexts.

Tone	CV	Gloss	CVʔ	Gloss	CVh	Gloss
/4/	βːe^{4}	*hair*	*	*	βeh^{4}	*beat.3sg (intr.)*
/3/	nːe^{3}	*plow*	nːeʔ^{3}	*straw rope*	nːeh^{3}	*dream*
/2/	nːe^{2}	*to lie (tr.)*	nːiʔ^{2}	*smelly*	βːeh^{2}	*cave*
/1/	nːe^{1}	*naked*	ʔniʔ^{1}	*be.salty*	cnãh^{1}	*brother (voc.)*
/45/	*	*	*	*	βːeh^{35}	*straw mat*
/13/	*	*	*	*	keh^{13}	*barely*
/43/	li^{43}	*small*	*	*	*	*
/32/	nːe^{32}	*water*	*	*	kʷːeh^{32}	*edible green*
/31/	nːe^{31}	*meat*	*	*	*	*

Itunyoso Trique has four level tones and five contour tones. Following normal IPA conventions, 5 is high and 1 is low here. Table 1 shows that seven of these tones occur in open syllables, seven in syllables with a coda /h/, and three in syllables with a coda /ʔ/. In context, coda /h/ is phonetically realized as [ɦ] or as vocalic breathiness on the preceding vowel, [a̤]. This is true even when it precedes a voiceless onset of a following word. Coda /ʔ/, by contrast, is normally realized with complete glottal closure in all contexts. Coda laryngeal consonants are both lexically and morphologically contrastive.

Due to the fact that more contrasts occur in final syllables in the language, the present study focuses on the coarticulation of tone only in final syllables in the different glottalization contexts in Table 1. Additionally, a variety of tones occur in the context of an intervocalic glottal stop, shown in Table 2. Intervocalic glottal stops are typically realized as creaky phonation without full glottal closure. As a comparison to the context of final glottalization, the realization of tone was examined in this intervocalic context. Spectrograms and waveforms demonstrating the realization of intervocalic and final glottalization are given in Fig. 2. Both of these examples were extracted from similar carrier sentences.

2. Methods

Three comparisons were made which tested the magnitude of tone–laryngeal coarticulation in Trique words: /Vː/ vs. /Vh/, /Vː/ vs. /Vʔ/, and /Vh/ vs. /VʔVh/. For each of these comparisons, F_0 and phonation type were investigated by extracting F_0 and spectral tilt measures from vowels in non-laryngealized and laryngealized syllable conditions.

2.1. Speech materials

The data for this study consisted of 96 monosyllabic and disyllabic words contrasting for both tone and one of four laryngeal contexts: open syllables, syllables with a coda /ʔ/, syllables with a coda /h/, and syllables with an intervocalic /ʔ/ with a coda /h/ (VʔVh). The entire sequence of each syllable condition was measured. The tone–laryngeal combinations shown in Tables 1 and 2 were investigated in each of these condition. Between three and five words were examined per tone–laryngeal condition, e.g. 3–5 words with tone /3/ and a coda /h/. There were 38 words in the /V/ condition, 9 in the /Vʔ/ condition, 35 in the /Vh/ condition, and 14 in the /VʔVh/ condition. The differences in the number of

Table 2
Tones in context of intervocalic /ʔ/.

σ_2	/4/	/3/	/2/	/1/	/43/	/32/	/45/
σ_1							
/4/	jã^{4}ʔãh^4 *guitar*	jã^{4}ʔãh^3 *god*	×	×	ru^{4}k^{w}i^{4}ʔi^{43} *peach*	×	×
/3/	jã^{3}ʔã^4 *to be hot*	ja^{3}ʔa^3 *brush*	sã^{3}ʔãh^2 *money*	kã^{3}ʔã^1 *breath*	×	jã^{3}ʔã^{32} *light, fire*	jo^{4}ʔoh^5 *land, dirt*
/2/	×	jã^{2}ʔã^3 *to be hot (outside)*	ta^{2}ʔah^2 *half (adj.)*	×	×	×	×
/1/	×	kã^{1}ʔã^3 *four*	×	na^{1}ʔah^1 *shame*	×	×	×

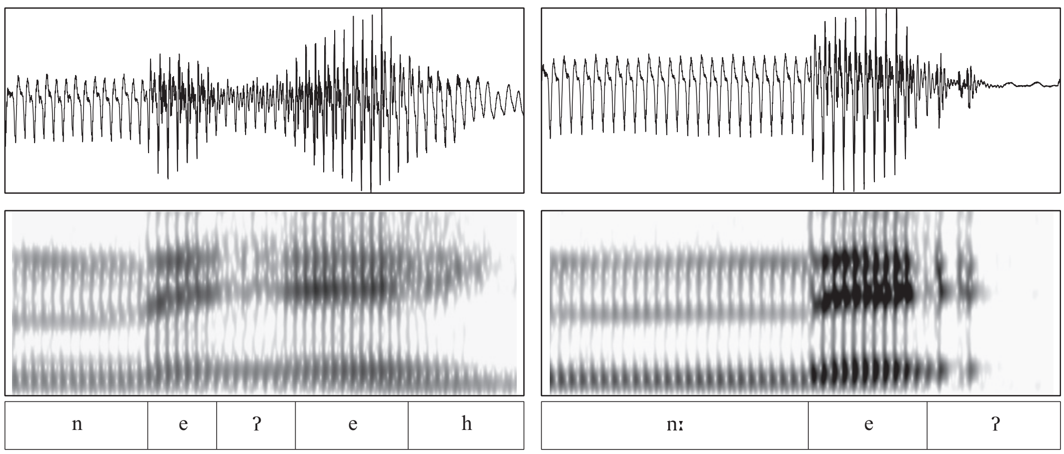

Fig. 2. Intervocalic /ʔ/ in word /ne^{3}ʔeh^3/ '*baby*' (left); Coda /ʔ/ in word /nːeʔ^3/ '*straw rope*' (right). Both words were extracted from carrier sentences. Long duration of glottal closure is typical of coda glottalization.

words for each phonological context partly reflect the way target words were collected during fieldwork, but are also a consequence of each context having fewer contrastive tones. Only three tones are contrastive before a coda /ʔ/, while 7 are contrastive in open syllables or before a coda /h/. Thus, due to the phonological restrictions on which tones are allowed to occur in different contexts, only the three level tones are comparable across contexts.

One stimulus had to be discarded from the /Vh/ syllable condition as not all speakers were familiar with the elicited word. This word was a specific lexical item /k^w:eh^3/ '*pus*'. Most speakers use the more general word /lo^{3}ʔloh^3/ '*mucus*.' Furthermore, upon further investigation, the elicited tokens in the tone /4/ + coda /ʔ/ context were mislabelled examples of tone /3/ + coda /ʔ/. Thus, no examples from the former context were examined.

Final syllable rimes from words in natural carrier sentences were examined. Given the nature of the fieldwork context, it was preferable to provide natural contexts for each of the words rather than unnatural contexts like "Say ___ again." Therefore, some of the carrier sentences differed. Most of the target words were nouns, which appeared in the carrier sentence /ni^{4}ʔja^{43}___ nã^3/, see.1sg ___ here/this, *I see* ___ *here*. The carrier sentence used for verbs or adjectives was /a^{3}kwa^{4}ni^{43}___ nã^3/, today ___ here/this, *Today* ___ *this*. While the carrier sentence used for adjectives and verbs was different from that which was used with nouns, the phonological context surrounding the target word was similar. In each case, the word following the target was identical and the tone preceding the target was identical, /43/, produced on an open syllable. The tones adjacent to the target were kept consistent to control for any tone–tone coarticulatory F_0 effects. Each carrier sentence was repeated 6 times for a total of 492 word repetitions per speaker. Sentences produced with disfluencies were discarded and not analyzed.

2.2. Speakers and data collection

Eight speakers were recruited for the investigations, four female, four male. Six speakers were between the ages of 18 and 26. One male speaker, (C), was 35 years old at the time of recording. Another female speaker, (G), was 56 years old. All participants were native, fluent speakers of Itunyoso Trique who were raised in San Martín Itunyoso. No participant reported having a history of speech or hearing disorders. Upon examining the data, two speakers recordings were eliminated (G and R). The recordings of these speakers had exceptionally low amplitude and contained substantial background noise (rain on a corrugated roof). Only six of the original eight speakers' data was used.

All speakers were native speakers of Trique with no history of speech or hearing disorders. All but one speaker were fully bilingual in Spanish. At the time of recording, there was no established orthography in the language.[3] Carrier sentences were read aloud by the author or by the author's main consultant as a verbal prompt to the speaker. The verbal prompt contained the target word in the carrier sentence. Speakers were asked to repeat this prompt five times. For seven of the speakers, recording took place in a quiet room in a house located in San Martín Itunyoso in Oaxaca, Mexico. The remaining speaker was recorded in his home in the central valley of California, USA.

[3] See DiCanio and Cruz Martínez (2010) for a recent development of the orthography.

Prior to recording, a consent form was presented to speakers in Spanish to read. For those speakers who were not literate, the consent form was translated and explained to them by the author's main consultant. After reading or hearing the details of the study, speakers supplied their verbal consent to participate in the acoustic investigation. Speakers who did not understand aspects of the investigation discussed their concerns with the author's primary consultant who acted as an interpreter.

Participants spoke into a uni-directional dynamic hand-held microphone that was maintained at a comfortable distance by the author. Recordings were made directly onto the author's Apple iBook G4 computer using an M-Audio MobilePre® USB preamplifier as an audio interface. Praat version 4.6 (Boersma & Weenink, 2009) was used to record all data. All data was sampled at 44.1 kHz.

2.3. *Measures*

In order to determine the timing relationships between F_0 perturbation and laryngeal activity, two measures of spectral tilt were examined along with F_0. Spectral tilt is a useful measure in examining the magnitude of glottal tension or spreading, and has been reliably used in the acoustic analysis of phonation type in a number of different languages (Blankenship, 2002; DiCanio, 2009; Esposito, 2006; Holmberg, Hillman, Perkell, Guiod, & Goldman, 1995; Kirk, Ladefoged, & Ladefoged, 1984; Kreiman, Gerratt, & Antoñanzas Barroso, 2007; Ladefoged, Maddieson, & Jackson, 1988; Pennington, 2005; Watkins, 2002). Two measures of spectral tilt were used: H1–H2 and H1–A3. The first measure usually distinguishes among breathy, modal, and tense or creaky phonation types. It reflects the difference in amplitude between the first harmonic, H1, and the second, H2. This corresponds to changes in the duration of the return phase of the glottal pulse (Kreiman et al., 2007). The second measure, H1–A3, has been reliably used to distinguish between breathy and non-breathy phonation types (Blankenship, 2002; DiCanio, 2009). It reflects the difference in amplitude between the first harmonic and the third formant (A3). Since, it corresponds to a larger frequency region, it is a better measure of global changes in the slope of the spectrum (Pennington, 2005). Furthermore, F3 varies less than other formants with changes in vowel quality; it is a more stable measure of spectral tilt differences than H1–A1 or H1–A2 are.

Individual sound files were first hand-labelled in associated textgrid files using Praat (Boersma & Weenink, 2009). For the /Vː/, /Vh/, and /VʔVh/ syllable conditions, vowels were measured from the release of the preceding onset consonant to the final glottal pulse on the rime target. Glottalization in the /VʔVh/ syllable condition was always realized as creaky phonation with no cessation of voicing. Thus, the duration of glottalization here was considered part of the analysis. For the /Vʔ/ syllable condition, glottal closure was abrupt, but often associated with a short duration of creaky phonation. All sound files were visually inspected in Praat and care was taken to exclude abrupt coda glottalization in the analysis window when Praat was unable to determine F_0 over this duration. When excessive glottalization made F_0 analysis impossible, these tokens were simply excluded from the analyzed data set. In the /Vh/ syllable condition, voiceless frication was rare. Instead, voicing persisted over the entire rime duration, where the coda was variably realized as vocalic breathiness or [ɦ]. Pitch was accurately estimated on such words. An example of the realization of this coda is shown in Fig. 3 for the word [βːeh[35]] '*straw mat*'.

Measures of spectral tilt and F_0 were taken using VoiceSauce, a Matlab script (Mathworks, 2009) developed by the Speech Processing and Auditory Perception laboratory at UCLA (Shue, Keating, & Vicenik, 2009). The program makes a number of different spectral analyses of labeled speech stimuli and has been used successfully for the analysis of non-modal phonation in various languages, such as Yi (Kuang, 2011) and Jalapa Mazatec (Garellek & Keating, 2011). For the analysis of spectral tilt, VoiceSauce estimates F_0 using the *Straight* method (Kawahara, de Cheveigné, & Patterson, 1998) and formant values using the Snack Sound Toolkit (Sjölander, 2004). Estimates of harmonic and formant amplitude are made F_0-synchronously over a four F_0 period window. All spectral tilt data in VoiceSauce are corrected to adjust for the effect of vowel formants, following the formula discussed in Iseli, Shue, and Alwan (2007). As such, all H1–H2 and the H1–A3 values reported here reflect this adjustment.

For the /Vʔ/ syllable condition, F_0, H1–H2, and H1–A3 measures were made at 4 even time indices along the duration of each vowel. For all other syllable conditions, these measures were made at 6 even time indices. This adjustment for the /Vʔ/ syllable condition was made to control for the shortened vowel duration on vowels preceding a coda /ʔ/. The average duration of vowels preceding a coda /ʔ/ was 113 ms, compared with 147 ms for vowels on open syllables and 166 ms for vowels in the /Vh/ syllable condition. For the calculation of F_0, the range was set to

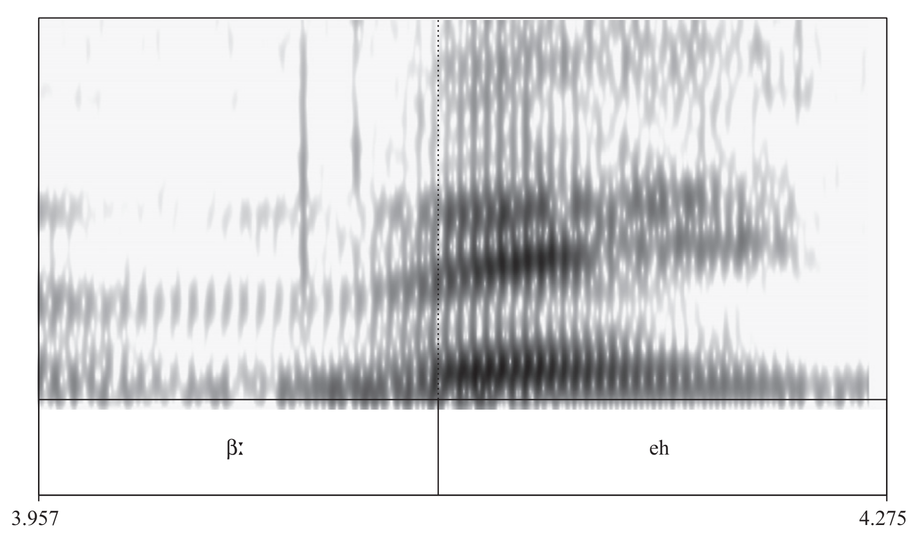

Fig. 3. Realization of Coda /h/ as [ɦ].

80–280 Hz. for male speakers and 100–450 Hz for female speakers. Following the calculation of these measures, data from all subjects were grouped together and statistically analyzed using R (R Development Core Team, 2009).

For each syllable condition, the results were normalized into *z*-scores and statistically analyzed using a generalized linear mixed effects model (lme4) with speaker treated as a random effect (Baayen, 2008; Bates, 2005). As dynamic measures were taken, time was treated as a discrete variable with six levels in this model. In mixed effects models, *p*-values are calculated not based on a standard degrees of freedom value, but on the upper bound of the degrees of freedom (*df=total observations−fixed effects parameters*). This is typical for mixed effects models, as the estimation of the degrees of freedom is not clearly established (Baayen, 2008; Bates, 2005). Two sets of *p*-values were obtained from the mixed effects model, one using Markov-chain Monte-Carlo (MCMC) sampling and another based on the *t* distribution. The *p*-values reported here derive from the *t* distribution, but were validated against those from the MCMC simulation, which adjusts for random effects. The value given with the *t*-statistic, e.g. *t*[num], reflects the upper bound on the degrees of freedom. This statistical analysis controls for the lack of balance in the current design and has been used in similar studies analyzing spectral tilt and F_0 in Jalapa Mazatec, a related language (Garellek & Keating, 2011).

3. Results

3.1. The effects of sex and F_0 on spectral tilt

In the following sections, three main linguistic factors: syllable condition, tone, and time, were included in the statistical models for each of the three acoustic measures (H1–H2, H1–A3, F_0). Separately, some differences were observed between the speakers' sexes, though this particular non-linguistic factor was not controlled for in the experimental design. It is well-established that female speakers have breathier voices than male speakers (Esposito, 2010; Garellek & Keating, 2011; Hanson & Chuang, 1999; Hanson, Stevens, Kuo, Chen, & Slifka, 2001; Klatt & Klatt, 1990). As such, females are generally expected to have steeper glottal source spectrums and higher spectral tilt values. As the posterior glottal opening is larger for females, more airflow is lost during glottal vibration. This venting of subglottal pressure impedes the magnitude of glottal closure, which results in less excitation of the higher frequencies in the source spectrum (Ladefoged et al., 1988).

Because it was not specifically controlled for, the effect of sex on the different acoustic measures was analyzed separately to determine whether it should be included as an additional statistical factor in the following sections. The effect of sex on each of the acoustic measures was analyzed in two different generalized linear mixed effects models. In the first analysis, *z*-score transformed F_0 values were analyzed in a model with sex and syllable condition treated as fixed effects and speaker treated as a random effect. In the second analysis, normalized H1–H2 and H1–A3 values were analyzed in a model with syllable condition, normalized F_0, and sex as fixed effects and speaker as a random effect.

For the F_0 measure, there was a significant main effect of sex ($t[2700]=5.4$, $p<.001$). Female speakers had higher F_0 values than male speakers. Data showing the pitch ranges for different speakers are given in Fig. 5.

Fig. 4 shows the effect of sex on spectral tilt values for each syllable condition. For the H1–H2 measure, there was a significant main effect of F_0 ($t[2700]=5.7$, $p<.001$). Higher F_0 values correlated with higher H1–H2 values. No main effect of sex on H1–H2 was found across syllable conditions, but there was a strong effect of sex on H1–H2 values in the /Vː/ syllable condition ($t[2700]=7.6$, $p<.001$). In addition, a strong $F_0 \times$ syllable condition effect was observed for this condition ($t[2700]=11.0$, $p<.001$). However, these two effects are not independent. Female speakers have higher F_0 ranges than male speakers do.

For the H1 − A3 measure, there was a significant main effect of F_0 ($t[2700]=4.1$, $p<.001$). Lower F_0 values correlated with higher H1–A3 values. No main effect of sex was found across syllable conditions, nor were there any significant interactions with sex. Significant $F_0 \times$ syllable condition interactions occurred for the /Vː/ syllable condition ($t[2700]=5.3$, $p<.001$) and for the /Vh/ syllable condition ($t[2700]=3.9$, $p<.05$). While the relationship between H1–A3 and F_0 was not as strong as the relationship between H1–H2 and F_0, the two spectral tilt measures showed opposite patterns. Higher F_0 values are correlated with higher H1–H2 values, while lower F_0 values are correlated with higher H1–A3 values.

Sex is a significant predictor of F_0 range. However, when it is included in models along with F_0, it is not significant. These findings suggest that sex-specific differences in the data are mostly attributable to variation in F_0. As such, Sex is not included in the statistical models for the spectral tilt measures in the following sections. It is included only in the analysis of F_0 differences.

3.2. /Vː/ vs. /Vh/

In each of Sections 3.2–3.4, general results from linear mixed effect models are presented first. Following this, interactions of general factors (syllable condition, time) with tone are presented. These interactions reflect tone-specific phonetic variability. While such findings are important in their own right, they are distinguished from the general findings for the sake of clarity.

3.2.1. Spectral tilt results

Fig. 6 shows the differences in spectral tilt for the /Vː/ and /Vh/ syllable conditions with tones /1/, /2/, /3/, /4/, and /32/. Each spectral tilt measure was analyzed using a generalized linear mixed effects model with three factors: syllable condition (/Vː/ vs. /Vh/), tone, and time. Speaker was treated as a random effect. Time was treated as a discrete variable in order to examine time-specific effects of spectral tilt on the speech signal. The main factor of syllable condition was significant ($t[10\,938]=2.0$, $p<.05$), but there were no significant interactions between syllable condition and time. This suggests that breathy phonation caused a general change in H1–H2 values, but none that were more strongly associated with specific portions of the rime. In general, the H1–H2 measure distinguished tones in the open syllable condition from the /Vh/ syllable condition, but the magnitude of this effect differed by tone.

Tones differed in how they were affected by the syllable condition. Significant interactions between tone and syllable condition were observed for tone 2 ($t[10\,938]=2.4$, $p<.05$), for tone 3 ($t[10\,938]=3.4$, $p<.001$), and for tone 4 ($t[10\,938]=2.4$, $p<.05$). No effect of syllable condition was observed for tone /32/. There was a significant effect of time on spectral tilt at time index 6 ($t[10\,938]=2.8$, $p<.01$). However, no significant interactions between syllable condition and time were observed. Tone /4/ was more substantially affected by changes in spectral tilt than the other tones were. This is reflected in a significant three-way interaction between tone 4, time index, and syllable condition ($t[10\,938]=3.0$, $p<.01$).

For the analysis of H1–A3, no general effect of syllable condition was observed, but there was a significant syllable condition × time

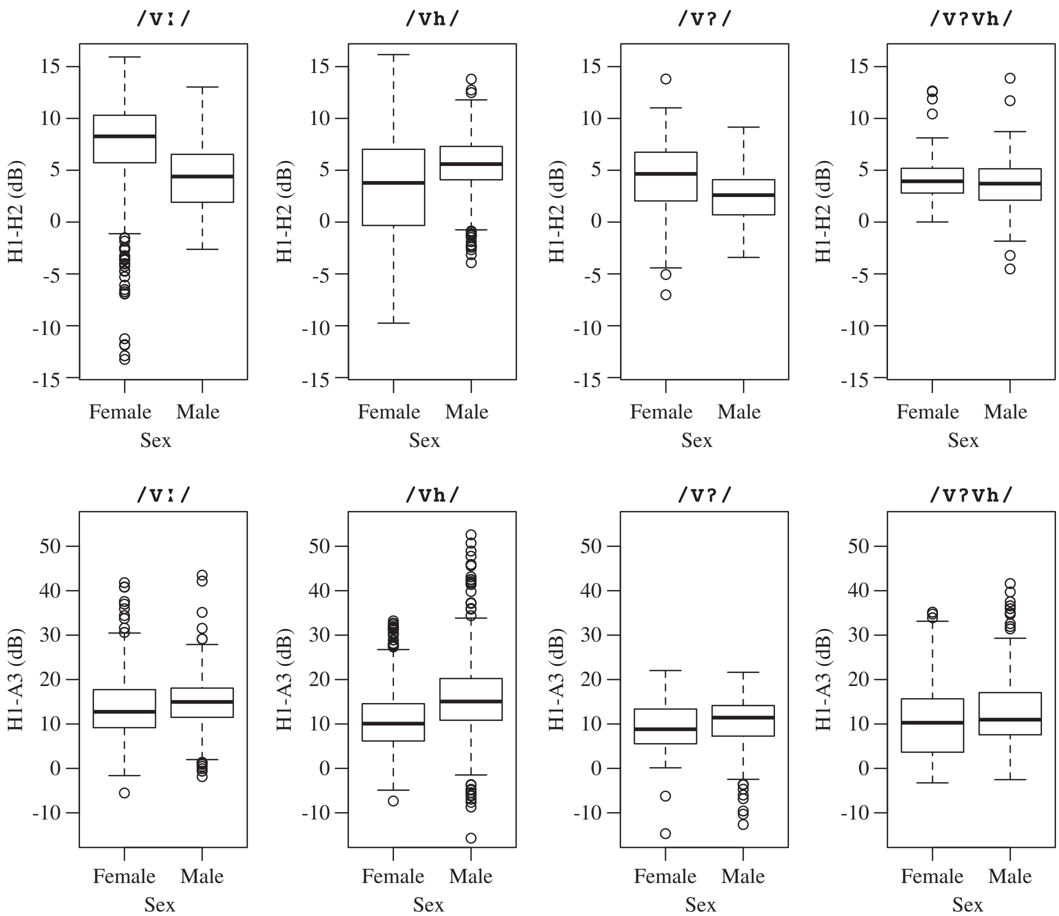

Fig. 4. Effect of sex on H1–H2 values, by syllable condition (top). Effect of sex on H1–A3 values, by syllable condition (bottom).

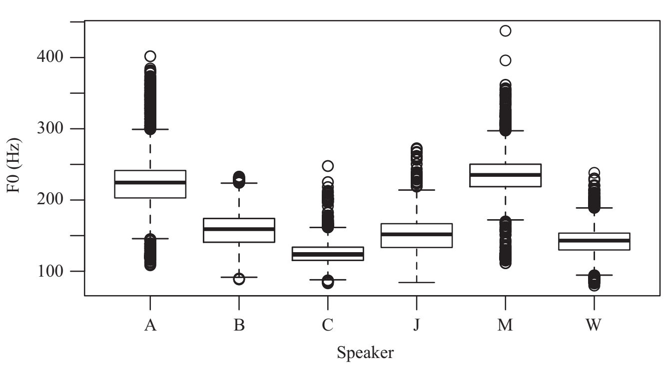

Fig. 5. Pitch range for different Trique speakers. Speakers A and M are female. All other speakers are male.

interaction at time index 6 ($t[10\,938]=2.4$, $p<.05$) and a significant general effect of time ($t[10\,938]=2.5$, $p<.05$). This finding reflects a lowering of H1–A3 at the end of the rime duration for tones in the /Vh/ syllable condition. A few tone-specific effects were observed as well. There was a significant time × syllable condition × tone interaction at time index 4 for tone /2/ ($t[10\,938]=3.2$, $p<.01$), tone /3/ ($t[10\,938]=2.9$, $p<.05$), and tone /32/ ($t[10\,938]=3.7$, $p<.001$). This finding was reflected in the significantly higher H1–A3 values for these tones at this time index.

These data demonstrate that breathy phonation in the /Vh/ syllable condition corresponds with changes in spectral tilt. However, the typical pattern is for vocalic breathiness to induce raising of H1–H2 values (Garellek & Keating, 2011; Ní Chasaide & Gobl, 1997; Pennington, 2005). In the data above, lower spectral tilt values are observed at the end of the rime duration in the context where breathy phonation is strongest. This finding is at odds with the previous literature. It is discussed in Section 4.2.

3.2.2. F_0 results

Fig. 7 shows the F_0 on tones on /Vh/ rimes and, as a comparison, on open syllables. In the /Vː/ condition, most of the phonological level tones are realized with no significant changes in F_0. However, tone /1/ falls slightly during the first half of the vowel duration. The F_0 differences between tones /1/ and /2/ are smaller than those between either tones /2/ and /3/ or /3/ and /4/.[4]

The F_0 measure was statistically analyzed using a generalized linear mixed effects model with four factors: syllable condition (/Vː/ vs. /Vh/), tone, time, and sex. Speaker was treated as a random effect and time was treated as a discrete variable. A significant main effect of syllable condition was found ($t[10\,938]=3.5$, $p<.001$) along with significant syllable condition × time interactions at time index 3 ($t[10\,938]=3.0$, $p<.01$), 4 ($t[10\,938]=2.9$, $p<.01$), 5 ($t[10\,938]=2.9$, $p<.01$), and 6 ($t[10\,938]=3.5$, $p<.001$). This finding reflects F_0 lowering associated with the presence of breathy phonation. A significant main effect of sex was also found in the data ($t[10\,938]=6.8$, $p<.001$), but with no significant interactions.

Non-modal phonation did not affect all tones equally. There were significant syllable condition × tone interactions for tone /2/ ($t[10\,938]=3.5$, $p<.001$), tone /3/ ($t[10\,938]=2.7$, $p<.01$),

[4] For an in-depth discussion of the phonetic differences between tones in open syllables, I refer the reader to DiCanio (2008).

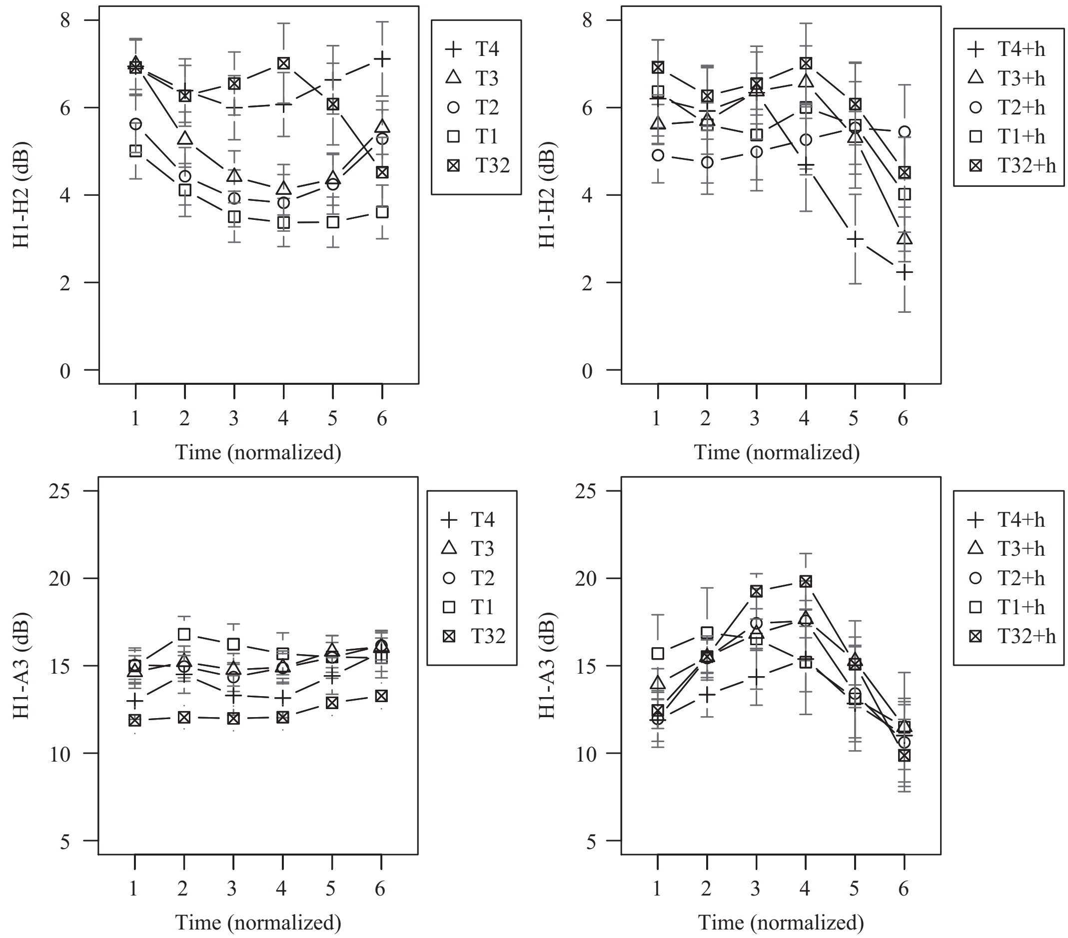

Fig. 6. Top: H1–H2 values for open syllable tones (left) and tones with coda breathy phonation (right). Bottom: H1–A3 values for open syllable tones (left) and tones with coda breathy phonation (right).

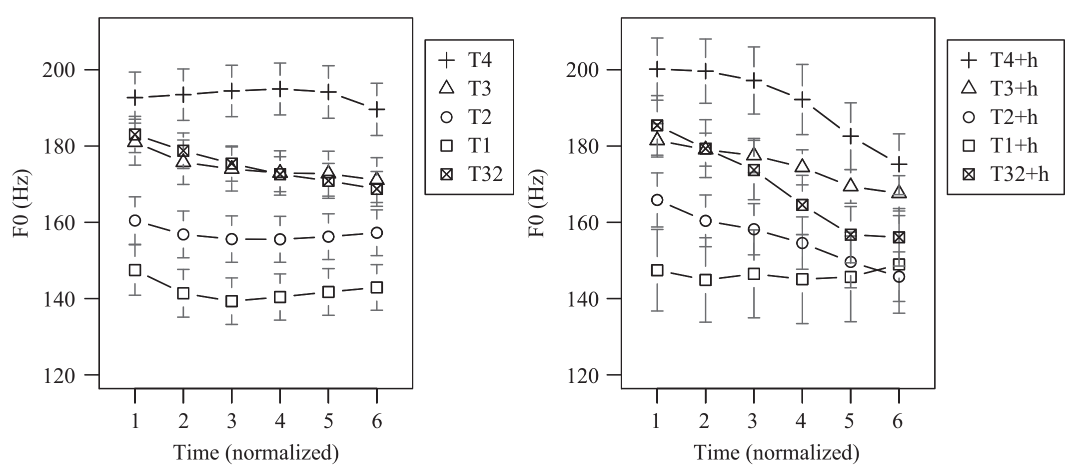

Fig. 7. F_0 on open syllables (Left) and on syllables with a Coda /h/ (Right).

tone /32/ ($t[10\,938]=4.8$, $p<.001$), and tone /4/ ($t[10\,938]=6.4$, $p<.001$). Surprisingly, a higher F_0 value was observed for tones /4/ and /2/ in the first third of the rime duration in the /Vh/ syllable condition than in the /Vː/ syllable condition. Furthermore, numerous significant three-way interactions were observed in the latter half of the /Vh/ rime. Certain tones underwent F_0-lowering earlier in this syllable condition than others. For tones /4/, /2/, and /32/, F_0 lowering begins after the first third of the rime duration. For tone /3/, F_0 lowering is only significant in the final third of the rime duration. The tones in the /Vh/ syllable condition are realized distinctively from the tones in the open syllable condition. Like the spectral tilt data in the previous section, the effects of the /Vh/ condition on F_0 are more significant in the final part of the /Vh/ rime.

The data here suggest a strong relationship between breathy phonation in the /Vh/ syllable condition and F_0 perturbation. Overall, breathy phonation causes an average F_0 fall of 10.1 Hz on the final half of the rime. For tones /4/ and /2/, the perturbation effect was stronger, as shown in Table 3. This finding is explainable due to the fact that these tones were produced with higher starting F_0 values in the /Vh/ syllable condition than in the /Vː/ condition.

3.2.3. Summary

The cross-condition comparison reveals that the presence of vocalic breathiness in the latter half of the rime duration in the /Vh/ syllable condition induces F_0 lowering effects. Such effects are absent in the /Vː/ syllable condition. The timing of breathiness in the /Vh/ syllable condition is gradual, not abrupt, but affects tones differently. Tone /1/ does not undergo F_0 lowering in the context of breathy phonation, while all other tones do. These findings support the hypothesis that significant F_0 effects occur in contexts where non-modal phonation type overlaps adjacent vowels.

3.3. /V/ vs. /V ʔ/

3.3.1. Spectral tilt results

In the /Vʔ/ syllable condition, a full [ʔ] is often produced with short duration creaky phonation on the preceding vowel. Thus, approximately half of the rime duration is characterized by substantial aperiodicity and/or full glottal closure (see Fig. 2). Since glottalization begins approximately halfway into the rime's duration, spectral tilt measures were only calculated over four time indices prior to glottalization. Since all data were normalized for time, this adjustment allows for a similar time:duration ratio across the /Vː/ and /Vʔ/ rimes. H1–H2 and H1–A3 values for /Vʔ/ rimes are shown in Fig. 8.

The spectral tilt measures were statistically analyzed using a generalized linear mixed effects model with three factors: syllable condition (/Vː/ vs. /Vʔ/), tone, and time. Speaker was treated as a random effect and time was treated as a discrete variable. For the H1–H2 measure, no general effect of syllable condition was observed. Though, there were significant tone-specific effects. For tone /2/, there was a significant syllable condition × tone interaction ($t[3480]=2.1$, $p<.05$). The H1–H2 trajectory of this tone differed from the others. Tone /3/ was realized with higher H1–H2 values than the other tones ($t[3480]=2.9$, $p<.01$) in both syllable conditions. No effect of time nor interactions with time were observed. In general, the two syllable conditions were not reliably distinguished by the H1–H2 measure.

For the analysis of the H1–A3 measure, each of the main effects was significant. A strong effect of syllable condition was observed ($t[3480]=4.6$, $p<.001$). All tones were realized with lower H1–A3 values in the /Vʔ/ syllable condition than in the /Vː/ condition. A significant syllable condition × time interaction occurred as well ($t[3408]=3.2$, $p<.01$). This reflects the change in H1–A3 values across the /Vʔ/ duration. With respect to tone-specific effects, tones /2/ and /3/ were realized with lower H1–A3 values at time index 1 while tone /1/ was realized with higher values. In general, the H1–A3 measure more strongly distinguished the two syllable conditions than the H1–H2 measure did.

Table 3
F_0 perturbation effect of breathy phonation, by tone.

T4 (Hz)	T3 (Hz)	T2 (Hz)	T1 (Hz)	T32 (Hz)
17.1	7.0	14.1	1.1	11.0

3.3.2. F_0 results

Fig. 9 shows the F_0 values for the three tones examined across these two syllable conditions.

The results from a generalized linear mixed effects model did not find any significant main effect of syllable condition on F_0. While tones in the /Vʔ/ syllable condition appear in Fig. 9 to have lower F_0 values than tones in the /Vː/ syllable condition, these differences were not significant. The lack of a general effect of condition on F_0 here may result from either the greater variability in the realization of F_0 prior to a coda /ʔ/ or the more abrupt timing (phasing) found with the coda. The first possibility is unlikely, however, as the standard deviation of F_0 prior to the coda /ʔ/ was smaller (31–39 Hz) than the standard deviation of F_0 in open syllables (40–42 Hz). The abruptness of the glottalization gesture may prevent significant coarticulatory F_0 effects. A significant effect of sex on F_0 was found ($t[3408]=9.1$, $p<.001$), but with no significant interactions.

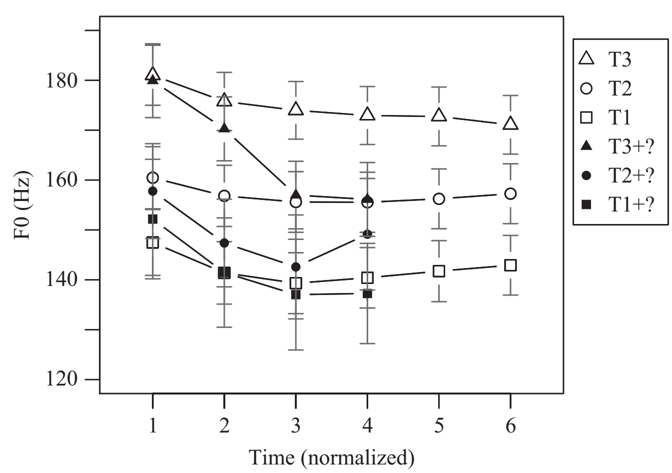

Fig. 9. F_0 values for tones in /Vː/ and /Vʔ/ conditions.

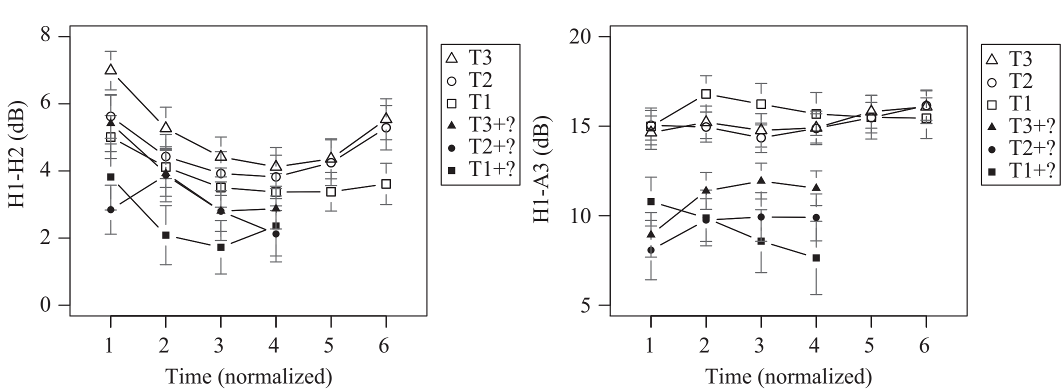

Fig. 8. H1–H2 values for tones in /Vː/ and /Vʔ/ syllable conditions (left); H1–A3 values for tones in /Vː/ and /Vʔ/ syllable conditions (right).

3.3.3. Summary

Within the /Vʔ/ syllable condition, lower H1–A3 values are observed throughout the vowel preceding glottalization. Lower spectral tilt values are anticipated before glottalization as a flatter source spectrum reflects increased tension between the vocal folds (Ladefoged et al., 1988; Pennington, 2005). However, H1–A3 was the only reliable spectral tilt measure distinguishing the two syllable conditions. The timing of glottalization in the /Vʔ/ syllable condition is abrupt. As a result, fewer coarticulatory changes in spectral tilt and F_0 were observed on the preceding vowel.

3.4. /Vh/ vs. /V ʔ Vh/

3.4.1. Spectral tilt results

In the /VʔVh/ syllable condition, a full [ʔ] is rarely produced. Instead, a very short modal vowel is produced, followed by a short period of aperiodicity, followed by a longer modal vowel which becomes breathy. This realization of glottalization is shown in Fig. 10. Such sequences are disyllabic in Itunyoso Trique, differing phonologically from true *interrupted* vowels in Oto-Manguean languages (Longacre, 1952; Silverman, 1997a), which are single vowel nuclei. As a comparison to the coda glottalization condition, intervocalic glottalization with coda /h/ was investigated in six tonal contexts in Itunyoso Trique: /1/, /3/, /4/, /32/, /35/, /13/. Tone /2/ does not surface in words with the shape /VʔVh/. Instead, this tone merges with tone /32/. This pattern is unsurprising, as tones /2/ and /32/ alternate throughout the morphology of Itunyoso Trique. Fig. 11 shows changes in H1–H2 and H1-A3 for level tones (left) and contour tones (right) for the /Vh/ and /VʔVh/ syllable conditions.

The spectral tilt measures were statistically analyzed using a generalized linear mixed effects model with three factors: syllable condition (/VʔVh/ vs. /Vh/), tone, and time. Speaker was treated as a random effect and time was treated as a discrete variable. A strong effect of syllable condition on H1–H2 was observed ($t[7710]=4.6$, $p<.001$). H1–H2 values were lower in the /VʔVh/ syllable condition than in the /Vh/ syllable condition. This corresponds to the presence of creaky phonation. Significant syllable condition × time interactions were observed in the latter third of the rime duration (time indices 5 and 6) ($t[7710]=5.0$, $p<.001$). This finding is reflected in lower H1–H2 values in the /VʔVh/ syllable condition, corresponding to the presence of breathy phonation.

With respect to tone-specific effects, a significant tone × syllable condition interaction was observed for tone /13/ ($t[7710]=2.2$, $p<.05$), tone /3/ ($t[7710]=2.2$, $p<.05$), tone /32/ ($t[7710]=2.8$, $p<.01$), and tone /4/ ($t[7710]=3.7$, $p<.001$). Tones /3/ and /32/ were more substantially affected by the presence of glottalization, while tones /13/ and /4/ were less so. In general, the H1–H2 measure strongly distinguished between tones produced in the context of creaky phonation and those produced in a modal context.

For the H1–A3 measure, a significant, but weaker effect of syllable condition was observed ($t[7710]=2.1$, $p<.05$). Tones produced in the /VʔVh/ syllable condition had lower H1–A3 values than tones produced in the /Vh/ syllable condition. A significant syllable condition × time interaction occurred at time index 5 ($t[7710]=2.1$, $p<.05$), reflecting a change from creaky phonation to modal voicing late in the rime duration in the /VʔVh/ syllable condition. Only one tone-specific effect was observed, for tone /3/ ($t[7710]=2.1$, $p<.05$). The most drastic shift in H1–A3 values occurred for this tone. Yet, in general, the H1–A3 measure less strongly distinguished between tones produced in creaky and modal phonation contexts.

In general, the H1–H2 and H1–A3 measures distinguished between the /VʔVh/ and /Vh/ syllable conditions. Though, the substantial variability between tones suggests that the production of intervocalic creak may be controlled in a specific way with respect to certain tones. For instance, insofar as the two measures of spectral tilt reflect glottal constriction, tone /13/ is produced with relatively little evidence of creaky phonation. Other tones are produced with altogether greater creaky phonation which is gradually phased across the duration of the adjacent vowels. Furthermore, contour tones are produced with greater final breathiness in the /VʔVh/ syllable condition than in the /Vh/ condition.

3.4.2. F_0 results

Fig. 12 displays F_0 measurements at six points across the duration of each syllable condition.

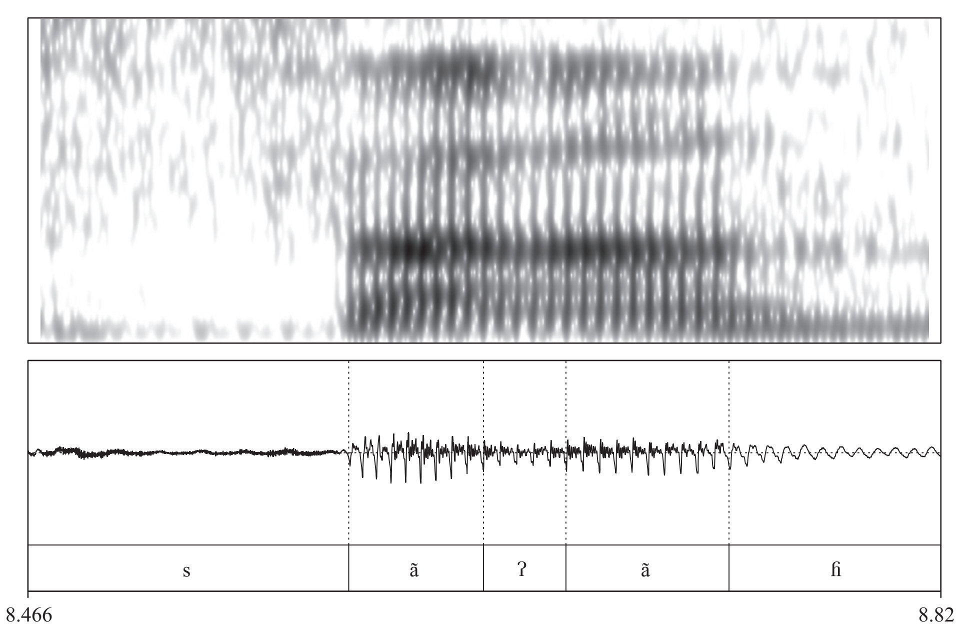

Fig. 10. Spectrogram and waveform of /sã³ʔãɦ²/ '*money*'.

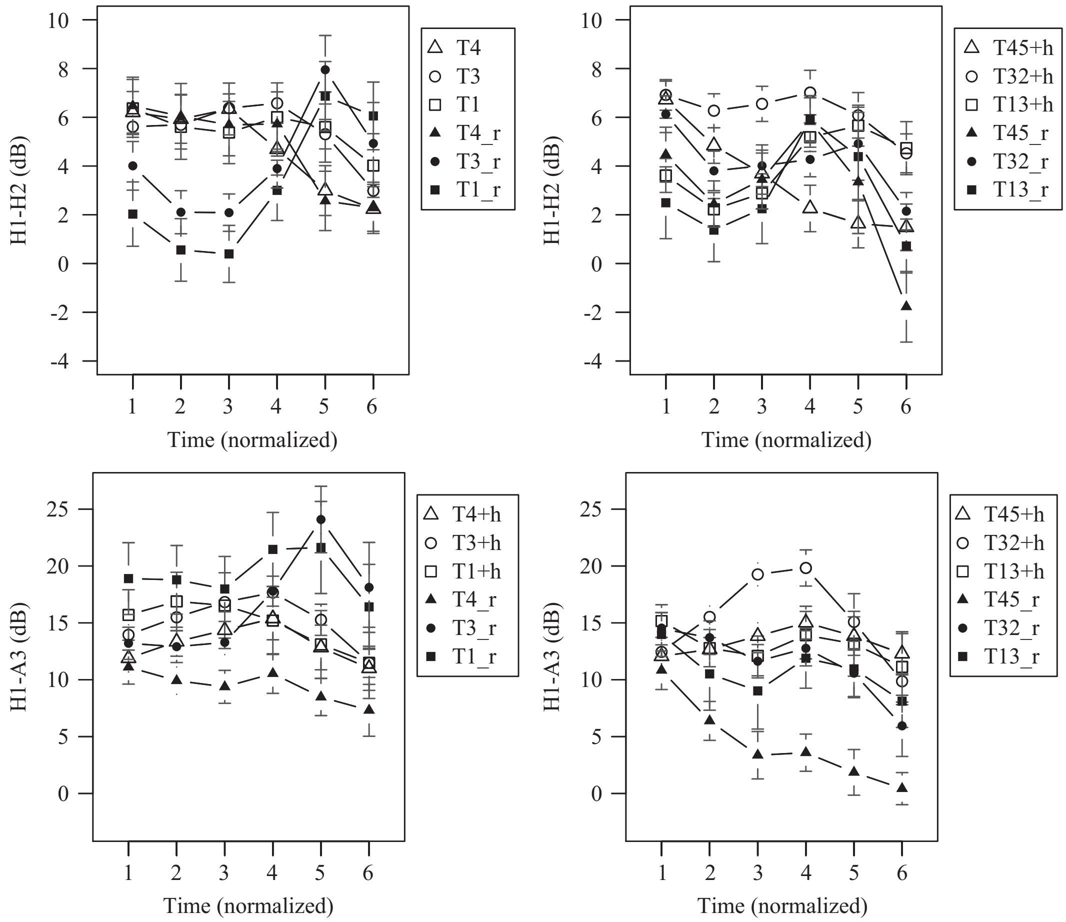

Fig. 11. Top: H1–H2 values for level tones in /Vh/ and /VʔVh/ conditions (left) and for contour tones in /Vh/ and /VʔVh/ conditions (right). Bottom: H1–A3 values for level tones in /Vh/ and /VʔVh/ conditions (left) and for contour tones in /Vh/ and /VʔVh/ conditions (right).

The F_0 measure was statistically analyzed using a generalized linear mixed effects model with four factors: syllable condition (/VʔVh/ vs. /Vh/), tone, time, and sex. Speaker was treated as a random effect and time was treated as a discrete variable. No significant effect of syllable condition was observed. A significant main effect of sex was also observed ($t[7710]=6.3$, $p<.001$), though with no significant interactions. A significant syllable condition × tone effect was found for tone /45/ ($t[7710]=3.6$, $p<.001$). This tone was realized with substantially lower F_0 values corresponding with the presence of creaky phonation. The data in Fig. 12 suggests similar F_0-lowering patterns for the other tones examined here, but their magnitude is visibly smaller than the pattern observed for tone /45/.

3.4.3. Summary

Intervocalic glottalization in the /VʔVh/ condition corresponds to lower H1–H2 and H1–A3 values on vowels in most tonal contexts. As observed by the strong tone × syllable condition interactions, individual tones were affected in distinct ways by glottalization. The presence of glottalization over tones /1/, /3/, /32/, and /45/ resulted in significant spectral tilt differences for these tones, but not for tones /13/ or /4/. Increased breathy phonation was observed in the /VʔVh/ syllable condition for contour tones in contrast with level tones. Despite the relatively strong effect of glottalization on spectral tilt, it had a much smaller effect on F_0. A significant F_0 perturbation effect was observed only for tone /45/. There is no general explanation for why this particular pattern should occur; it seems to result from language-specific tone–laryngeal coarticulation strategies.

4. Discussion and conclusion

4.1. Glottal timing strategy

There are three patterns observed in Itunyoso Trique data, corresponding to the three separate laryngealized syllable conditions. In the /Vh/ context, strong breathy phonation occurs, gradually phased from approximately 50% of the [Vh] rime to the end, over an average duration of 83 ms. In the /Vʔ/ context, an abrupt glottal stop occurs over the latter half of the rime duration, over an average duration of 113 ms. In the /VʔVh/ context, slight creak is gradually phased over an average duration of 60 ms, interrupting the surrounding vowels. After this, modal phonation returns but is then gradually phased with increasing breathiness. The abrupt phasing pattern of [ʔ] in the /Vʔ/ syllable condition differs from the gradual pattern in the /VʔVh/ syllable condition. Greater F_0 and greater spectral tilt differences were observed in the syllable conditions with more gradual glottal consonant phasing than in the /Vʔ/ syllable condition. Changes in spectral tilt on the vowel correlate with changes in F_0. The presence of breathy phonation is correlated with lower H1–H2 values and lower F_0. The presence of creaky phonation is correlated with lower H1–H2 values and lower F_0 as well. This finding confirms the hypothesis stated in Section 1.2: the magnitude of coarticulatory overlap of

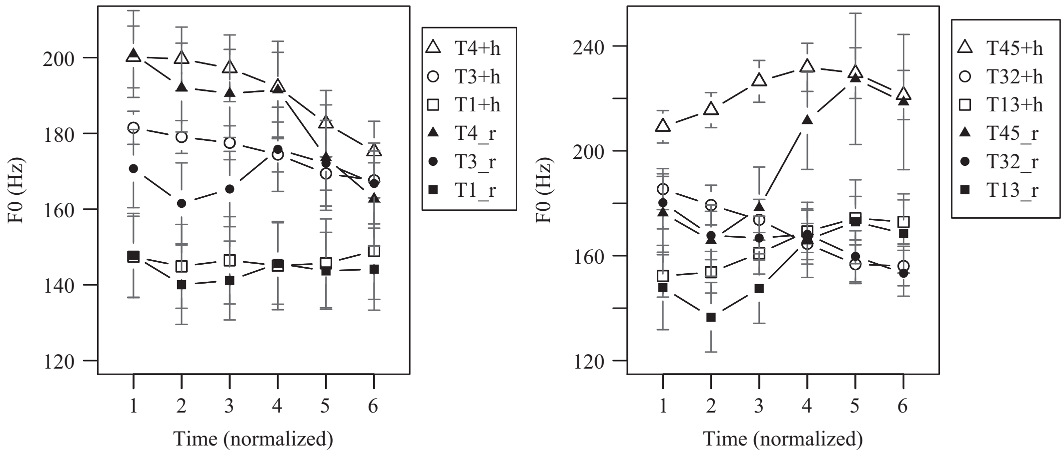

Fig. 12. F_0 values for level tones in /Vh/ and /VʔVh/ conditions (left) and for contour tones in /Vh/ and /VʔVh/ conditions (right).

non-modal phonation type on the vowel correlates with the magnitude of F_0 lowering.

Both the shorter duration of the intervocalic glottalization and the fact that the coda [ʔ] is delimitative may contribute to the observed phasing differences. Decreased duration is often a source for the loss of complete closure in oral stops (Lavoie, 2001), via a process of *target undershoot* (Flemming, 2001; Kirov & Gafos, 2007; Lindblom, 1983). Undershoot of complete glottal closure has been discussed in a related language, Coatzospan Mixtec, where intervocalic glottalization is often realized with only slight F_0 and amplitude perturbations and no noticeable creak (Gerfen & Baker, 2005; Hillenbrand & Houde, 1996).[5] A similar process is described for intervocalic /h/ in English (Pierrehumbert & Talkin, 1992). Lenited intervocalic glottalization is also a characteristic of certain Mixtecan languages spoken near the Puebla-Oaxaca border (in Tepejillo Mixtec, Micaltepec Mixtec, Cacaloxtepec Mixtec, and Ixtapan Mixtec) (Josserand, 1983, p. 274). In a few of these languages, glottalization may be completely lost. It is conceivable that a similar process of glottal lenition is occurring in Itunyoso Trique.

However, all laryngeal codas are equally delimitative in Trique, as they are the only permitted coda segments. Without a more general explanation, the differences in phasing between coda /h/ and /ʔ/ in Itunyoso Trique must result from language-specific glottal timing strategies associated with particular glottal segments. In a similar way, Jun (1995) argues, for Korean, that lenis stop voicing is sensitive to prosodic structure. The glottal adduction gesture in lenis stops occurs as a byproduct of the closure duration, which varies according to the prosodic constituent in which it appears. Yet, the glottal abduction gesture in fortis stops in Korean is insensitive to durational differences resulting from prosodic structure. It is conceivable that a similar distinction between glottal timing patterns occurs in Itunyoso Trique with coda /h/ and coda /ʔ/. Two distinct consonants show different patterns of intergestural timing in the same prosodic position. Laryngeal phasing patterns may simply be language-specific phonetic patterns independent from perceptual constraints on phonetic implementation.

4.2. *F_0 perturbation effects*

Three F_0 perturbation patterns, corresponding to the three laryngealized syllable conditions, were observed in the Trique data. In the /Vh/ syllable condition, F_0 lowering was observed for all tones but tone /1/. In the /Vʔ/ syllable condition, no F_0 effect was observed. In the /VʔVh/ syllable condition, intervocalic glottalization caused significant F_0 perturbations only for tones /3/ and /45/. For the syllable conditions with gradually phased non-modal phonation, F_0 lowering occurred. Perturbation occurs most robustly in those cases where the laryngealization is longer in duration. Where more non-modal phonation occurs across the vowel duration, there is more substantial F_0 perturbation.

The Trique data demonstrate that non-modal phonation type may have specific effects on certain tones but none on others. In the /Vh/ and /VʔVh/ syllable conditions, no F_0 effect was observed for tone /1/, even though significant changes in spectral tilt were observed for this tone in the latter condition. This pattern may have a mechanical explanation. Since Itunyoso Trique has four level tones (and a fifth phonetic level for rising tones), the F_0 target for tone /1/ may naturally approach a lower limit in the speaker's F_0 range. The lack of F_0 perturbation effects on this tone may be due to this natural F_0 "floor." Despite changes in the magnitude of glottalization, very low tones cannot be lowered any further. A separate tone effect was observed in the /VʔVh/ syllable condition. Tone /45/ was more strongly affected by the presence of intervocalic glottalization than the other tones were. Further work is need to understand the significance of this pattern.

The influence of F_0 on H1–H2 values reported in Section 3.1 provides an explanation for the lower H1–H2 values observed in the /Vh/ syllable condition. Generally speaking, one predicts higher H1–H2 values in the presence of breathy phonation. One predicts *lower* H1–H2 values with lower F_0 values (Garellek & Keating, 2011). It is generally understood that breathy phonation causes F_0 to lower. Does this associated F_0 lowering cause lowering or raising of H1–H2? Both possibilities are predicted from the findings in the current literature. Yet, breathy phonation induces F_0-lowering in the Trique data. This finding suggests that changes in H1–H2 as a result of breathy phonation in the /Vh/ syllable condition were more strongly associated with the associated F_0 perturbation effects than with changes in phonation type. This explanation diverges from the one given for creaky phonation type, where the direction of change in spectral tilt correlates with the direction of F_0 change. Finally, it is worth noting that differences in the durational properties for individual tones and differences in speech rate may have influenced the degree to which tones were influenced by coarticulatory effects of non-modal phonation. The influence of tonal duration on coarticulation is a promising topic for future research.

[5] It is possible that spectral tilt differences accompanied glottalization in these languages too, though these authors did not investigate this.

Within the literature on tonogenesis, languages with glottal consonants evolve into systems with contrastive tone. The coarticulatory patterns observed here suggest a path by which such patterns begin. Greater coarticulatory non-modal phonation induces pitch changes on a portion of the rime. While these changes are coincidental with changes in phonation type, pitch cues begin to be used by listeners in their identification of the phonation contrast. In some languages, such as Tamang, evidence of this process is observed in the compensation between pitch and phonation type by speakers (Michaud & Mazaudon, 2006). Where lower pitch values occur, less breathy phonation occurs. Where higher pitch values occur, more breathy phonation occurs. In the /Vh/ syllable condition here, significant pitch lowering occurs with the presence of breathy phonation. This lowering pattern may have historically influenced tonal patterns in other Trique variants. For instance, in Chicahuaxtla Trique (Longacre, 1959), the tonal pattern cognate with Itunyoso Trique tone /45/ is /453/. In Itunyoso Trique, tone /45/ always surfaces with a coda /h/. However, this particular tone in Chicahuaxtla Trique never surfaces with a coda /h/.[6] It is likely that the lower final pitch (/3/) in the Chicahuaxtla Trique tonal contour resulted from coarticulatory changes in pitch associated with the presence of a historical coda /h/.

4.3. Conclusions

The Itunyoso Trique data demonstrate that both abrupt and gradual phasing occur in the production of glottal consonants. An increase in coarticulatory overlap between non-modal phonation and adjacent vowels results in greater coarticulatory F_0 perturbation effects. Abrupt glottal gestures do not correlate with significant changes in F_0. Moreover, not all tones are equally affected by non-modal phonation. Low tones do not undergo F_0 perturbation regardless of the magnitude of observed coarticulatory changes in voice quality. These effects suggest a close relationship between the magnitude of coarticulatory overlap in phonation type and the magnitude of coarticulatory pitch perturbation effects.

Acknowledgments

Acknowledgments are given to Benigno Cruz Martínez and the community of San Martín Itunyoso. This paper also benefitted from helpful comments by researchers at the Acoustical Society of America Annual Meeting in San Antonio, Texas, 2009, Ken deJong, Egidio Marsico, Ioana Chitoran, and four anonymous reviewers.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.wocn.2011.10.006.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.
Bates, D. M. (2005). Fitting linear mixed models in R. *R News*, *5*, 27–30.
Beddor, S. P., Harnsberger, J. D., & Lindemann, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates. *Journal of Phonetics*, *30*, 591–627.
Bird, S., Caldecott, M., Campbell, F., Gick, B., & Shaw, P. A. (2008). Oral-laryngeal timing in glottalised resonants. *Journal of Phonetics*, *36*, 492–507.
Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, *30*, 163–191.

[6] This is even an exception to a regular morphological pattern in the language involving coda /h/ insertion (DiCanio, 2010a).

Boersma, P., Weenink, D. (2009). Praat: Doing phonetics by computer [computer program]. 〈www.praat.org〉.
Browman, C., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, *3*, 219–252.
Browman, C., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In: J. Kingston, & M. E. Beckman (Eds.), *Papers in laboratory phonology 1: Between the grammar and the physics of speech* (pp. 341–376). Cambridge University Press.
Brunelle, M. (2003). Coarticulation effects in Northern Vietnamese tones. Unpublished manuscript.
Byrd, D. (1996a). Influence on articulatory timing in consonant sequences. *Journal of Phonetics*, *24*, 209–244.
Byrd, D. (1996b). A phase window framework for articulatory timing. *Phonology*, *13*, 139–169.
Chávez Peón, M. E. (2010). *The interaction of metrical structure, tone, and phonation types in Quiaviní Zapotec*. Ph.D. thesis. The University of British Columbia.
Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In: C. Gussenhoven, & N. Warner (Eds.), *Papers in laboratory phonology*, Vol. 7). Berlin, New York: Mouton de Gruyter.
Cho, T., Jun, S.-A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*, 193–228.
Dart, S. (1987). An aerodynamic study of Korean stop consonants: Measurements and modeling. *Journal of the Acoustical Society of America*, *81*(1), 138–147.
DeJong, K., & McDonough, J. (1993). Tone and tonogenesis in Navajo. UCLA Working Papers in Phonetics: Fieldwork Studies of Targeted Languages (Vol. 84, pp. 165–182).
DiCanio, C. (2010a). The structure of the Itunyoso Trique Word. manuscript.
DiCanio, C. T. (2008). *The phonetics and phonology of San Martín Itunyoso Trique*. Ph.D. thesis. University of California, Berkeley.
DiCanio, C. T. (2009). The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association*, *39*(2), 162–188.
DiCanio, C. T. (2010). Illustrations of the IPA: San Martín Itunyoso Trique. *Journal of the International Phonetic Association*, *40*(2), 227–238.
DiCanio, C. T., & Cruz Martínez, B. (2010). *Chungwì Snáhánj nìh, El Mundo Triqui: Palabras de San Martín Itunyoso*. Instituto Nacional de Lenguas Indígenas.
Diffloth, G. (1989). Proto-Austroasiatic creaky voice. *Mon-Khmer Studies*, *15*, 139–154.
Dürr, M. (1987). A preliminary reconstruction of the Proto-Mixtec Tonal system. *Indiana: Contributions to the Ethnology and Archaeology, Linguistics, Social Anthropology, and History of Indigenous Latin America*, *11*, 19–60.
Esposito, C. (2004). Santa Ana del Valle Zapotec phonation. UCLA Working Papers in Phonetics (Vol. 103, pp. 71–105).
Esposito, C. (2006). *The effects of linguistic experience on the perception of phonation*. Ph.D. thesis, UCLA.
Esposito, C. (2010). Variation in contrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association*, *40*, 181–198.
Farnetani, E., & Recasens, D. (1999). Coarticulation models in recent speech production theories. In: W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data, and techniques* (pp. 31–65). Cambridge University Press. (Chapter 2).
Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, *18*, 7–44.
Francis, A. L., Ciocca, V., Ka Man Wong, V., & Ka Lam Chan, J. (2006). Is fundamental frequency a cue to aspiration in initial stops?. *Journal of the Acoustical Society of America*, *120*(5), 2884–2895.
Frazier, M. (2009). *The production and perception of pitch and glottalization in Yucatec Maya*. Ph.D. thesis. University of North Carolina, Chapel Hill.
Garellek, M., & Keating, P. (2011). The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association*, *41*(2), 185–205.
Gathercole, G. (1983). Tonogenesis and the Kickapoo Tonal system. *International Journal of American Linguistics*, *49*(1), 72–76.
Gerfen, C., & Baker, K. (2005). The production and perception of laryngealized vowels in Coatzospan Mixtec. *Journal of Phonetics*, *33*, 311–334.
Gerratt, B. R., & Kreiman, J. (2001). Toward a taxonomy of nonmodal phonation. *Journal of Phonetics*, *29*, 365–381.
Gibbon, F., Hardcastle, W. J., & Nicolaidis, K. (1993). Temporal and spatial aspects of lingual coarticulation in /kl/ sequences: A cross-linguistic investigation. *Language and Speech*, *36*, 261–277.
Gick, B. (1999). A gesture-based account of intrusive consonants in English. *Phonology*, *16*, 29–54.
Gick, B., Campbell, F., Oh, S., & Tamburri-Watt, L. (2006). Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. *Journal of Phonetics*, *34*, 49–72.
Good, C. (1979). *Diccionario Triqui. Serie de Vocabularios Indigenas*, Vol. 20. Mexico: Summer Institute of Linguistics.
Gordon, M., & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, *29*, 383–406.
Han, J.-I. (2007). The role of vowel contrast in language-specific patterns of vowel-to-vowel coarticulation: Evidence from Korean and Japanese. In: *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 509–512).
Hanson, H. M., & Chuang, E. S. (1999). Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Journal of the Acoustical Society of America*, *106*, 1064–1077.
Hanson, H. M., Stevens, K. N., Kuo, H.-K. J., Chen, M. Y., & Slifka, J. (2001). Towards models of phonation. *Journal of Phonetics*, *29*, 451–480.

Hardcastle, W. J. (1973). Some observations on the tense-lax distinction in initial stops in Korean. *Journal of Phonetics*, *1*, 263–272.

Haudricourt, A. G. (1954). De l'origine de tons en Viênamien. *Journal Asiatique*, *242*, 69–82.

Hillenbrand, J. M., & Houde, R. A. (1996). Role of F_0 and amplitude in the perception of intervocalic glottal stops. *Journal of Speech and Hearing Research*, *39*, 1182–1190.

Hollenbach, B. E. (1977). Phonetic vs. phonemic correspondence in two Trique dialects. In: W. R. Merrifield (Ed.), *Studies in Otomanguean phonology. Publications in Linguistics* (Vol. 54, pp. 35–67). Dallas: Summer Institute of Linguistics.

Hollenbach, B. E. (1984a). Copala Trique tone and universal features. *Coyote Papers*, *5*, 96–119.

Hollenbach, B. E. (1984b). *The Phonology and morphology of Tone and Laryngeals in Copala Trique*. Ph.D. thesis. University of Arizona.

Hollenbach, B. E. (2004). *Gramática popular del triqui de Copala*. Mexico: Summer Institute of Linguistics.

Hollenbach, B. E. (2007). *Vocabulario breve del triqui de San Juan Copala*. SIL International.

Holmberg, E., Hillman, R., Perkell, J., Guiod, P., & Goldman, S. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language, and Hearing Research*, *38*, 1212–1223.

Hombert, J. M. (1975). Perception of contour tones: An experimental investigation. In: *Proceedings of the 1st annual meeting of the Berkeley Linguistics Society*, Vol. 1.

Hombert, J. M. (1976). Development of tone from segmentals: Evidence from contour tone perception. In: *Proceedings of the 8th international congress of the phonetic sciences*, Leeds.

Hombert, J. M. (1979). Consonant types, vowel height and tone. In: V. A. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 77–111). New York: Academic Press.

Hombert, J. M., Ohala, J. J., & Ewan, W. (1979). Phonetic explanations for the development of tones. *Language*, *55*(1), 37–58.

Howe, D., & Pulleyblank, D. (2001). Patterns and timing of glottalisation. *Phonology*, *18*, 45–80.

Hu, F., & Xiong, Z. (2010). Lhasa tones. In: *Proceedings of the 5th international conference on speech prosody*.

Huffman, M. K. (2005). Segmental and prosodic effects on coda glottalization. *Journal of Phonetics*, *33*, 335–362.

Iseli, M., Shue, Y.-L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America*, *121*(4), 2283–2295.

Josserand, J. K. (1983). *Mixtec dialect history*. Ph.D. thesis. Tulane University.

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody*. Ph.D. thesis. The Ohio State University.

Jun, S.-A. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In: B. Connell, & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in laboratory phonology*, Vol. 4 (pp. 235–253). Cambridge University Press. (Chapter 17).

Kawahara, H., de Cheveigné, A., & Patterson, R. D. (1998). An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: Revised Tempo in the straight suite. In: *Proceedings of the 5th international conference on spoken language processing*. .

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and overlap in the speech chain. *Language*, *82*(1), 33–63.

Kingston, J. (1985). *The phonetics and phonology of the timing of oral and glottal events*. Ph.D. thesis. University of California, Berkeley.

Kingston, J. (2005). The phonetics of Athabaskan tonogenesis. In: S. Hargus, & K. Rice (Eds.), *Athabaskan prosody*. Amsterdam: John Benjamins.

Kingston, J., Diehl, R. L., Kirk, C. J., & Castleman, W. A. (2008). On the internal perceptual structure of distinctive features: the [voice] contrast. *Journal of Phonetics*, *36*, 28–54.

Kirk, P. L., Ladefoged, P., & Ladefoged, J. (1984). Using a spectrograph for measures of phonation types in a natural language. UCLA Working Papers in Phonetics (Vol. 59, pp. 102–113).

Kirov, C., & Gafos, A. (2007). Dynamic phonetic detail in lexical representations. In: J. Trouvain, & W. J. Barry (Eds.), *Proceedings of the 16th international congress of phonetic sciences* (pp. 637–640). University of Saarbrücken.

Klatt, D., & Klatt, L. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*, 820–857.

Kochetov, A. (2002). *Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization*. New York, London: Routledge.

Kochetov, A. (2006). Syllable position effects and gestural organization: Articulatory evidence from Russian. In: L. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Papers in laboratory phonology. Varieties of phonological competence*, Vol. 8. Berlin, New York: Mouton de Gruyter.

Krakow, R. A. (1989). *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures*. Ph.D. thesis. Yale University.

Krakow, R. A. (1993). Nonsegmental influences on velum movement patterns: Syllables, sentences, stress, and speaking rate. In: M. A. Hufman, & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum* (pp. 87–116). New York: Academic Press.

Krakow, R. A. (1999). Physiological organization of syllables: A review. *Journal of Phonetics*, *27*, 23–54.

Kreiman, J., Gerratt, B., & Antoñanzas Barroso, N. (2007). Measures of the glottal source spectrum. *Journal of Speech, Language, and Hearing Research*, *50*, 595–610.

Kuang, J. (2011). *Production and Perception of the Phonation Contrast in Yi*. Master's thesis. UCLA.

Ladefoged, P., & Maddieson, I. (1985). Tense and lax in four minority languages of China. *Journal of Phonetics*, *13*, 433–454.

Ladefoged, P., Maddieson, I., & Jackson, M. (1988). Investigating phonation types in different languages. In: O. Fujimura (Ed.), *Vocal physiology: Voice production, mechanisms and functions* (pp. 297–317). New York: Raven Press, Ltd.

Lavoie, L. M. (2001). *Consonant strength: Phonological patterns and phonetic manifestations. Outstanding dissertations in linguistics*. Garland Publishing, Inc.

Lee, S. J. (2008). *Consonant-tone interaction in optimality theory*. Ph.D. thesis. Rutgers: The State University of New Jersey.

Lindblom, B. (1983). The economy of speech gestures. In: P. F. MacNeilage (Ed.), *The production of speech* (pp. 217–243). Springer-Verlag.

Longacre, R. E. (1952). Five phonemic pitch levels in Trique. *Acta Linguistica*, *7*, 62–81.

Longacre, R. E. (1959). Trique tone morphemics. *Anthropological Linguistics*, *1*(4), 5–42.

Maddieson, I. (2004). Timing and alignment: A case study of Lai. *Language and Linguistics*, *5*(4), 729–755.

Maddieson, I., & Emmorey, K. (1985). The relationship between semivowels and vowels: Cross-linguistic investigations of acoustic difference and coarticulation. *Phonetica*, *42*(4), 163–174.

Manuel, S. Y. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America*, *88*(3), 1286–1298.

Manuel, S. Y. (1999). Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. In: W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data, and Techniques* (pp. 179–198). Cambridge University Press. (Chapter 8).

Mathworks (2009). *Matlab: The Language of Technical Computing (2009). Version 7.9.0.529*. Natick, MA: Mathworks, Inc.

Matisoff, J. (1973). Tonogenesis in Southeast Asia. In: L. Hyman (Ed.), *Consonant types and tone: Southern California occasional papers in linguistics*, Vol. 1 (pp. 73–95). University of Southern California.

Mazaudon, M., & Michaud, A. (2008). Tonal contrasts and initial consonants: A case study of Tamang, a 'missing link' in tonogenesis. *Phonetica*, *65*(4), 231–256.

Michaud, A., & Mazaudon, M. (2006). Pitch and voice quality characteristics of the lexical word-tones of Tamang, as compared with level tones (Naxi data) and pitch-plus-voice-quality tones (Vietnamese data). In: *Speech prosody* (pp. 823–826).

Miller, A. L. (2007). Guttural vowels and guttural co-articulation in Ju|'hoansi. *Journal of Phonetics*, *35*, 56–84.

Munhall, K. G., & Löfqvist, A. (1992). Gestural aggregation in speech: Laryngeal gestures. *Journal of Phonetics*, *20*, 93–110.

Myers, S. (2000). Boundary disputes: The distinction between phonetic and phonological sound patterns. In: N. Burton-Roberts, P. Carr, & G. J. Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues*. Oxford: Oxford University Press.

Ní Chasaide, A., & Gobl, C. (1997). Voice source variation. In: W. J. Hardcastle, & J. Laver (Eds.), *The handbook of phonetic sciences*. Blackwell.

Öhman, S. E. G. (1967). Numerical model of coarticulation. *Journal of the Acoustical Society of America*, *41*, 310–320.

Peng, S.-h. (1997). Production and perception of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics*, *25*, 371–400.

Pennington, M. (2005). *The phonetics and phonology of glottal manner features*. Ph.D. thesis. Indiana University.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In: *Papers in laboratory phonology: Gesture, segment, prosody* (Vol. 2, pp. 90–117). Cambridge University Press.

Przezdziecki, M. A. (2005). *Vowel harmony and coarticulation in three dialects of Yoruba: Phonetics determining phonology*. Ph.D. thesis. Cornell University.

R Development Core Team, Vienna, A. (2009). R: A language and environment for statistical computing [computer program] 〈http://www.R-project.org〉, R Foundation for Statistical Computing.

Scarborough, R. A. (2004). *Coarticulation and the Structure of the Lexicon*. Ph.D. thesis. UCLA.

Shue, Y.-L., Keating, P., & Vicenik, C. (2009). VOICESAUCE: A program for voice analysis [computer program]. *Journal of the Acoustical Society of America*, *126*(2221(A)).

Silverman, D. (1997a). Laryngeal complexity in Otomanguean vowels. *Phonology*, *14*, 235–261.

Silverman, D. (1997b). *Phasing and recoverability. Outstanding dissertations in linguistics*. Routledge.

Silverman, D., & Jun, J. (1994). Aerodynamic evidence for articulatory overlap in Korean. *Phonetica*, *51*, 210–220.

Sjölander, K. (2004). *The snack sound toolkit computer program*. KTH, Sweden: TMH, Speech, Music and Hearing.

Svantesson, J.-O., & House, D. (2006). Tone production, tone perception, and Kammu tonogenesis. *Phonology*, *23*, 309–333.

Thongkum, T. (1991). An instrumental study of Chong registers. In: J. Davidson (Ed.), *Essays in Mon-Khmer linguistics in honor of H.L. Shorto* (pp. 141–160). London: School of Oriental and African Studies.

Thurgood, G. (2002). Vietnamese and tonogenesis. *Diachronica*, *19*(2), 333–363.

Watkins, J. (2002). *The phonetics of Wa: Experimental phonetics, phonology, orthography, and sociolinguistics*, Vol. 531. Pacific Linguistics.

Whalen, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics*, *18*(1), 3–35.
Wright, R. (1996). *Consonant clusters and cue preservation in Tsou*. Ph.D. thesis. UCLA.
Wright, R., Hargus, S., & Davis, K. (2002). On the categorization of ejectives: Data from Witsuwit'en. *Journal of the International Phonetic Association*, *32*(1), 43–77.
Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of the Acoustical Society of America*, *96*(4), 2240–2253.
Xu, C., & Xu, Y. (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association*, *33*(2), 165–181.
Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*, 61–83.
Xu, Y. (1998). Consistency of tone-syllable alignment across different syllable structures and speaking rates. *Phonetica*, *55*, 179–203.
Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, *27*, 55–105.
Xu, Y. (2004). Understanding tone from the perspective of production and perception. *Language and Linguistics*, *5*, 757–797.
Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication*, *46*, 220–251.
Xu, Y., & Liu, F. (2006). Tonal alignment, syllable structure, and coarticulation: Toward an integrated model. *Italian Journal of Linguistics*, *18*, 125–159.
Xu, Y., & Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America*, *111*, 1399–1413.
Zee, E. (1980). The effect of aspiration on the F_0 of the following vowel in cantonese. UCLA Working Papers in Phonetics (Vol. 49, pp. 90–97).
Zsiga, E., & Nitisaroj, R. (2007). Tone features, tone perception, and peak alignment in Thai. *Language and Speech*, *50*, 343–383.