UCLA Working Papers in Phonetics, No.111 (2012), pp. 92-122

Word-initial glottalization and voice quality strengthening*

Marc Garellek
(marcgarellek@ucla.edu)

Abstract

Despite abundant research on the distribution of word-initial glottal stops, it is still unclear which factors matter most in predicting where glottal stops occur and why. In this study, logistic mixed-effects regression modeling is used to predict the occurrence of word-initial full glottal stops ([ʔ]) in an English corpus. The results indicate that prominence and phrasing are overwhelmingly the most important factors in predicting full glottal stop occurrence. Moreover, prominent word-initial vowels that are not preceded by [ʔ] show acoustic correlates of glottal constriction, whereas non-prominent phrase-initial vowels do not. Rather, phrase-initial voicing (even for sonorants) is less regular, but in a manner inconsistent with glottal constriction. These findings are subsequently confirmed using articulatory measures from electroglottography, and extended to Spanish. Based on the results, a prominence-driven theory of word-initial glottalization is proposed and motivated, with higher phrasal domains responsible for the strength of the glottal stop gesture.

1 Introduction

1.1 Optional word-initial glottalization: a near-universal property of language?

Glottal stops occur before vowel-initial words (e.g., English 'apple' pronounced as [ʔæpl]) in many languages. In some (e.g., Arabic, Ilokano) glottal stops are thought to obligatorily mark word-initial vowels, whereas in others, like English, the presence of word-initial glottal stops is more variable (Lombardi 2002, Borroff 2007, Hayes 2009). This variable phenomenon, which I refer to here as 'word-initial glottalization,' might occur in all languages except those that contrast /#ʔV/ and /#V/. For example, word-initial glottalization is banned in Tongan, where words like /aa/ 'heat sticks over fire' and /ʔaa/ 'awake' contrast. It is indeed remarkable that word-initial glottalization is so common a phenomenon: no other segment is used epenthetically in the same environment and in so many languages. The goal of this study is to understand why glottal stops are typologically common before vowel-initial words.

A theory of word-initial glottalization should ultimately account for the phenomenon's cross-language prevalence. But for such a theory to be valid, its underlying assumptions need validation. For example, in my earlier statement, ('no other segment is used epenthetically in the same environment and in so many languages'), I assumed that glottal stops are segments (equivalent to, e.g, [t]), and that they are inserted, rather than arising through fortition of preexisting segments or

*I would like to thank Gabe Langton, Sam Slama, James Park, Karen Villanueva, Sara Haakana, and Yoo Kyung Kang for help with data collection and annotation, and Dulce Montoya and Victoria Mateu for help with devising the materials for the Spanish experiment and subject recruitment. I would also like to thank Pat Keating, Abeer Alwan, Sun-Ah Jun, Jody Kreiman, Megha Sundara, and Stefanie Shattuck-Hufnagel for their comments and discussion. Results from the corpus experiment were presented at the 2012 LabPhon meeting in Stuttgart, and an earlier analysis can be found in Garellek (2012). This work was supported in part by a doctoral fellowship to the author from the *Fonds québécois de recherche sur la société et la culture* (grant number 138016).

features, as I explain later. Further, I assumed that their reason of occurrence before vowel-initial words is the same across languages. These assumptions are proposed, if not adopted, in many studies of glottalization (Pierrehumbert & Talkin 1992, Dilley, Shattuck-Hufnagel & Ostendorf 1996, Borroff 2007), but are they justified? For example, if we treat glottal stops as segments, we must ask why researchers find that they are usually (even in prosodically-strong environments) realized non-canonically, as a form of irregular voicing or laryngealization (Pierrehumbert & Talkin 1992, Dilley et al. 1996). Why call them glottal 'stops' if they are rarely plosives? For this reason, the term 'glottalization' is often used to refer to both voicing irregularity (irregular pitch periods) and glottal stops (Redi & Shattuck-Hufnagel 2001). This in turn leads to another assumption, namely that all instances of glottalization can be treated as reflexes of a single phenomenon.

Even if not all instances of voicing irregularity are derived from the same phenomenon, it is clear that word-initial glottalization at least sometimes results in an incomplete glottal stop ([ʔ̰], a form of laryngealization), and this too has implications for theories of word-initial glottalization. If glottal stops are epenthetic, then we must reconcile the fact that the speaker goes through the added effort to epenthesize, while failing to realize this inserted segment in its canonical form.

In sum, previous attempts to characterize word-initial glottalization have made important theoretical assumptions that have yet to be validated: is word-initial glottalization to be treated as epenthesis? If so, of a segment or a feature? What should we count as a token of word-initial glottalization? In order to develop a valid theory of word-initial glottalization, I believe we should first determine what should count as an instance of glottalization, which in turn will allow us to determine the most important factors in predicting when the phenomenon occurs. Finally, we can then determine why word-initial glottalization occurs, and if it can be treated equally across languages, and if so, why it is so typologically common.

1.2 When does word-initial glottalization occur?

Although cross-linguistically widespread, it is clear that the frequency of occurrence of word-initial glottalization may differ across languages. For example, it is thought to be rare in Spanish, common in English and German, and almost across-the-board in Czech (Bissiri & Volín 2010, Bissiri, Lecumberri, Cooke & Volín 2011, Pompino-Marschall & Żygis 2011). Many researchers have investigated (for a variety of languages) the factors that promote the occurrence of word-initial glottalization. (Here I discuss only word-initial glottalization. For factors that affect coda glottalization or glottal replacement, see Milroy, Milroy, Hartley & Walshaw (1994), Pierrehumbert (1995), Huffman (2005) and Eddington & Channer (2010), among others.) Predictors of word-initial glottalization may be segmental, lexical, prosodic, or sociolinguistic. In English, segmental factors include hiatus (V#V) environments (Umeda 1978, Dilley et al. 1996, Pierrehumbert 1995, Mompeán & Gómez 2011, Davidson & Erker 2012) and word-initial back vowels are found to glottalize more frequently than non-back vowels (Umeda 1978). As for lexical factors, content words exhibit more frequent glottalization than function words (Umeda 1978). Women are known to use glottalization more than men (Byrd 1994, Dilley et al. 1996). Most studies emphasize the role of prosody, with presence of stress and/or a pitch accent on the word-initial vowel or later in the word, as well as a larger juncture with the preceding word, as factors (Pierrehumbert & Talkin 1992, Pierrehumbert 1995, Dilley et al. 1996). In other languages, additional factors that promote the occurrence of word-initial glottalization include word length in Dutch (Jongenburger & van Heuven 1991), presence of a preceding pause (Kohler 1994) as well as speech rate (all of which are correlated with other prosodic factors), and low vowel quality (Brunner & Żygis 2011, Pompino-Marschall & Żygis 2011) for German.

Despite abundant interest in the topic, there has been little investigation of which factors are most important for promoting word-initial glottalization. That is, we know what factors play a role in the occurrence of word-initial glottalization, but we do not know how important they are compared to each other. Again, some factors are correlated with others, e.g. vowel hiatus and pitch accent (Redi & Shattuck-Hufnagel 2001), changes in speech rate are correlated with changes in prosody, and the distinction between function vs. content words correlates with differences in lexical frequency. A model predicting the occurrence of word-initial glottalization should take into account all these factors, and determine which are most important, so that we can understand why it occurs.

1.3 Why is word-initial glottalization so common?

If, as mentioned in the preceding section, prosody plays an important role in determining when word-initial glottalization occurs, it is likely that the phenomenon is (sometimes) a form of prosodic strengthening of word-initial vowels (Fougeron 2001). Prosodic strengthening is the process by which articulations are 'strengthened' in prosodically strong environments, notably phrase-initially and under prominence (Keating 2006). By 'strengthening,' the articulation itself can become more forceful (Fougeron 2001), or the contrast between the target and neighboring sounds can be enhanced (Hsu & Jun 1998, Cho 2005). The former is usually called 'paradigmatic enhancement,' in contrast to the latter, which is 'syntagmatic.'

If word-initial glottalization is a form of prosodic strengthening, it is unclear what in fact is being strengthened. Is voice quality generally more forceful in strong environments? More forcefully-articulated voicing should yield laryngealization, which has increased glottal closure (Gordon & Ladefoged 2001), and there is some evidence that the start of an English utterance has tenser voice quality compared to later portions (Epstein 2002). Strengthening of voice quality provides a straightforward account of why word-initial glottalization is cross-linguistically common, because a more forceful articulation of voicing should result in laryngealization in every language. But if voicing in general is strengthened, then two important facts should be considered. First, in terms of theory, glottal stops should then not be regarded as distinct segments, but as the extreme result of voicing strengthening, as discussed by Pierrehumbert & Talkin (1992), Dilley et al. (1996), and Borroff (2007), among others. Thus, an incomplete glottal stop should not be regarded as a form of lenition, because it is stronger than the default, which is modal voicing. Rather, a full [ʔ] is the extreme case of this fortition (see discussion of similar issues regarding English /t/ affrication by Buizza & Plug (2012)). Second, in terms of theory prediction, if word-initial glottalization is a form of voice quality strengthening, then all voiced sounds, including voiced consonants, are expected to show increased laryngealization or even glottal stops in strong environments. However, if vowels are strengthened differently than other voiced sounds, then word-initial glottalization (by means of glottal stop insertion) can be specific to word-initial vowels. Under this view, a glottal stop might be obligatory or optional before all vowel-initial words, and lenited to modal voicing in the weakest prosodic environments. Note that it might be impossible to determine phonetically whether word-initial glottalization is obligatory or optional, if we posit that lenited glottal stops can be realized as modal voicing in the weakest environments. However, if glottal stops are always present (but lenited in weak environments), then this has problematic implications for typology and phonological analyses: it would then be unclear why word-initial glottalization is typologically common, and why 'optional' word-initial glottal stops do not pattern with obligatory glottal stops (Borroff 2007).

1.4 The current study

The goals of this paper are to determine (1) whether word-initial glottalization is largely a prosodic phenomenon, and if so, (2) whether the phenomenon arises from strengthening of voice quality or from glottal stop insertion that is specific to word-initial vowels. I address these goals in two studies. In Section 2, I analyze a corpus of American English to determine where full glottal stops ([ʔ]) occur. I focus on full glottal stops because other cases of irregular voicing might be due to other sources (Garellek 2012). In the corpus study, I also determine whether the same environments show acoustic signs of laryngealization (with increased vocal fold closure) when no glottal stop is present. In Section 3, I use a combination of acoustic and articulatory analyses to confirm the acoustic results from the first study. In Section 4, I synthesize the findings from these two studies, and propose a revised prosodic account of word-initial glottalization.

2 Corpus analysis of word-initial glottal stops

In this study, I investigate which factors best predict word-initial glottalization, to determine whether the phenomenon is truly prosodic. I also use acoustic analyses to determine if voice quality is generally strengthened in prosodically-strong conditions.

2.1 Method

The corpus analyzed here is the Boston University (BU) Radio News Corpus (Ostendorf, Price & Shattuck-Hufnagel 1995). The main motivation for using the BU radio news corpus was the fact that it is labeled for prosody. Another reason was that it has been analyzed for glottalization, both vowel-initial and in all word positions, by Dilley et al. (1996) and Redi & Shattuck-Hufnagel (2001). Thus, comparison with previous work is facilitated by using the same corpus. The section of the corpus used in the present work is from the Labnews corpus, consisting of radio news read in the laboratory. The four speakers analyzed in this study form a subset of the newscasters analyzed by Dilley et al. (1996). All speakers read the same news reports. The speakers were adults aged 25 to 40 years old, and with no perceived regional accent. A subset of the available corpus, two female speakers (f1a, f2b) and two male speakers (m1b, m2b), is analyzed below.

In the Tones and Break Indices (ToBI) prosodic annotation of the corpus, the break index tier is used to mark the prosodic grouping of words in an utterance. The end of each word is coded for the perceived strength of its association with the next word, on a scale from 0 (for the strongest perceived juncture) to 4 (for the most disjoint). A break index of 3 usually corresponds to the end of an intermediate phrase (ip) in English, whereas a break index of 4 typically corresponds to the end of an intonation phrase (IP). Because in this corpus the presence of a breath or pause following a break index of 4 was transcribed, I will refer to this as the end of an utterance within a breath group, labeled as break index 5 (cf. Price, Ostendorf, Shattuck-Hufnagel & Fong 1991). The utterance domain above the phrasal one has been shown to exhibit greater levels of prosodic strengthening (Fougeron & Keating 1997, Keating, Cho, Fougeron & Hsu 2003), and thus could be relevant for the present study.

In the present study, all vowel-initial words were extracted from the corpus. A total of 2010 vowel-initial words were extracted for the four speakers. For further details, see Garellek (2012). In addition to word-initial vowels, 1298 word-initial sonorants ([j, w, l, ɹ, m, n]) were extracted, as well as the following vowels. For example, for a word like 'Massachusetts,' the initial [m] and following [æ] were extracted from the corpus. Sonorant-initial words will be used to determine whether word-initial glottalization is found for other voiced sounds. Voiced obstruents were not included here due to frequent devoicing in conversational English. In total, 1291 vowels following word-initial sonorants were extracted. This means that seven post-sonorant vowels were not extracted from the total of 1298 sonorant-initial words. These were all cases of a sonorant followed by a syllabic [ɫ̩] (e.g. *will* pronounced as [wɫ̩]), where the boundary between vowel and coda was hard to determine or did not exist.

The corpus was then coded by two undergraduate research assistants trained in labeling acoustic irregularity. The coders were unaware of the purpose of the study, and thus were unbiased in their coding. In addition to coding for presence and type of irregularity (to be described below), the coders also transcribed the presence of a variety of other factors, described in further detail below. The agreement rate for codings of irregularity was over 90%. I then reviewed the corpus data and arbitrated between-coder differences.

Generally, the coding of voicing irregularity followed that described by Dilley et al. (1996) and Redi & Shattuck-Hufnagel (2001). First, the coders rated whether there was a percept of 'glottalization', regardless of whether the percept was due to a glottal stop. Tokens with weak percepts of glottalization with no clear acoustic evidence for voicing irregularity were labeled as glottalized, unlike in Dilley et al. (1996), where such tokens were excluded from the analysis. We included these tokens for the purposes of the quantitative analysis described below. Such tokens represented only 2% of total words in the corpus, and therefore were unlikely to have a significant influence on the subsequent analysis.

Second, if there was a percept of glottalization, the coders labeled the 'type' of aperiodicity found, based on inspection of the waveform. This labeling provides visual support for the percept of glottalization, but the individual types of aperiodicity will not be analyzed below. Four types were identified, three of which (aperiodicity, diplophonia, and creak) following the description by Redi & Shattuck-Hufnagel (2001). In the present study, a full glottal stop [ʔ] was also identified. For details on criteria used for labeling glottal stops and other types of aperiodicity, as well as their relative frequencies in the corpus, see Garellek (2012).

2.1.1 Other factors in the analysis

In addition to coding for presence and type of glottalization, the coders also recorded prosodic, lexical, and segmental information. The prosodic factors are summarized in Table 1. The factor 'prominence' refers to a syllable with prosodic prominence, either due to the presence of a pitch accent, already transcribed, or if the syllable belonged to a function word, but had an unreduced vowel (e.g. [ænd] for 'and'), or both. Thus, 'prominence' here represents a superset of pitch-accented syllables. The reason for including this factor was that some vowels had perceived prominence, but no pitch accent was marked in the BU Corpus. The absence of expected vowel reduction is not typically used as a cue for vowel prominence, though it is used as a cue for some degree of lexical stress (Hayes 1995). Further, the coders agreed that these words were more prominent than expected. Often these words occurred phrase-initially, suggesting that the absence of vowel reduction is related to phrase-initial strengthening (Cho & Keating 2009). Thus, in phrase-initial position, vowels – even when not pitch-accented – are perhaps nonetheless more strongly articulated and thus more perceptually prominent than they would be phrase-medially.

Aside from prosodic factors, lexical and segmental factors were also included, and they are summarized in Table 2. Lexical frequencies were taken from the $\text{SUBTLEX}_{\text{WF}}$-US corpus (Brysbaert & New 2009).

Table 1: Prosodic analysis factors.

Factor	Explanation
Preceding break index	Break index (from 0-5) between target word and prec. word
Following break index	Break index (from 0-5) between target word and foll. word
Pitch accent	Presence of pitch accent on target syllable
Pitch accent type	Type of pitch accent on target syllable (H*, L*, etc.)
Prominence	Presence of a pitch accent and/or unreduced, stressed vowel
Boundary tone	Presence of boundary tone/phrase accent on target syllable
Boundary tone type	Type of boundary tone/phrase accent on target syllable
Preceding pause	Presence of a pause before target syllable
Preceding glot.	Presence of glottalization-like irregularity on preceding syllable

Table 2: Segmental and lexical factors in the analysis.

Factor	Explanation
Vowel	The type of vowel in the target syllable
Vowel height	Whether the target vowel was high, mid, or low
Vowel frontness	Whether the target vowel was front, central, or back
Vowel length	Whether the target vowel was tense or lax
Word	The word containing the target syllable
Log frequency of word	Log frequency of target word
Word type	Whether target word a content or function word
Preceding sound	Final sound of preceding word
Hiatus	Potential for hiatus (i.e. prec. sound was a vowel)
Vowel quality of prec. vowel	Height, frontness, and length of prec. vowel.
Prec. word	The word preceding the target syllable
Log freq. of prec. word	Log frequency of the prec. word
Prec. word type	Whether prec. word was content or function word

2.1.2 Acoustic measures

In addition to the qualitative factors mentioned above, I obtained quantitative data from acoustic measures to provide a gradient analysis of the strength of glottalization. The acoustic measure included in the analysis here is H1*-H2*, obtained from Voice Sauce (Shue, Keating, Vicenik & Yu 2011), though results from additional measures are described in Garellek (2012). Voicing with increased closure typically shows lower values of H1*-H2* (see discussion in Garellek & Keating (2011)), and lower values of the measure (either corrected or uncorrected for vowel formants) are correlated with higher values of flow adduction quotient (Holmberg, Hillman, Perkell, Guiod & Goldman 1995), increased values of EGG contact quotient (DiCanio 2009, Kuang 2011, Esposito 2012), and lower open quotient derived from glottal area (Shue, Chen & Alwan 2010). Therefore, H1*-H2* is taken to be the likeliest acoustic measure to be correlated with increased glottal closure.

To obtain the measures, the coders manually verified the segment boundaries of the word-initial vowels in the corpus. Although segment boundaries had already been provided in the corpus, many had been aligned automatically, and many files had not been checked for segment boundaries. The boundaries of nasals and laterals were taken to be at the change in amplitude from the sonorant to the vowel. For [j], the F2 maximum was used as the starting boundary, and the beginning of the F2 fall as the end boundary. For [w], the boundaries were marked at the points where F3 was weak or not visible. For diphthongs like [oʊ], the entire post-sonorant diphthong was extracted. VoiceSauce was then run over the entire sound file (not just the labeled vowels), because many tokens were so short that they required longer windows of analysis in order to obtain acoustic measures. The values obtained for all data fames within a vowel token were then averaged over the entire vowel's duration, giving one value per token.

2.2 Results

Of the 2010 vowel-initial words extracted from the corpus, 1060 or 53% showed at least one form of irregularity. Only 300 or about 15% of all word-initial vowels had full glottal stops. Vowel-initial words showing a form of voicing irregularity with no glottal stop accounted for 37% of all vowel-initial words and about 72% of cases of irregular word-initial vowels (vowels with a glottal stop, aperiodicity, diplophonia, and/or creak). None of the vowels after sonorants (e.g., the [æ] in 'Massachusetts') had glottal stops, but about 20% showed voicing irregularity. 30% of the sonorants had irregular voicing, but only three cases of glottal stops before sonorants were documented. This number is virtually negligible, and these cases are likely instances of creak with a long lag between the first pulse and the next. Across speakers, the rates of voicing irregularity and glottal stops were 27% of all tokens for f1a, 39% for f2b, 26% for m1b, and 31% for m2b. Further details of the distribution of glottal stops in the corpus can be found in Garellek (2012).

2.2.1 Predicting full glottal stop occurrence

To predict where full glottal stops [ʔ] occur, the data were first subset into cases of word-initial vowels with a full [ʔ], and cases of no perceived/visual voicing irregularity. A mixed-effects logistic regression model was fitted to these data using the *lmer()* function in the *lme4* package (Bates, Maechler & Dai 2008) in *R* (R Development Core Team 2011), following Baayen (2008). The model's dependent variable was presence of [ʔ] vs. no perceived/visual voicing irregularity, and had 14 independent variables from the factors listed in Tables 1 and 2 above: previous break, pitch accent, prominence, hiatus, following break, word type, vowel height, length, and backness, presence of preceding pause and glottalization, word frequency, preceding word frequency, and preceding word type. An interaction term (presence of hiatus:preceding glottalization) was included because it improved the model's fit (which was assessed by the *anova* function in *R*, following Baayen (2008)). Speaker and word were included as random intercepts. The results are shown in Table 3. The coefficient estimates indicate the direction of significance, with a positive coefficient indicating an increase in the odds of there being a full glottal stop. Both an increase in preceding break index and presence of prominence on the following vowel increased the likelihood of glottal stop occurrence, and these factors were the most significant in the model. The effects of phrasal domain and prominence can be seen in Figure 1. Prominent vowels were more likely to be preceded by [ʔ], regardless of the preceding break. But the phrasal domain was also significant, with rates of glottal stop occurrence decreasing with a decrease in preceding break index.

Other significant predictors included presence of a preceding pause and preceding 'glottalization,' both of which increased the likelihood of obtaining a full [ʔ]. A preceding pause might increase the likelihood of a glottal stop for two reasons. First, pauses had already been marked in the corpus, but it was apparent to the coders that some of these represented the closure durations for glottal stops rather than true pauses. Second, true pauses increase the dissociation between two words, such that a break index of 4 with no pause is weaker than a 4 followed by a pause. Preceding glottalization (mostly from phrase-final creak) might increase the likelihood of there being a following glottal stop because the vocal folds are mostly abducted and closing irregularly during creak (Slifka 2006). Thus, vocal fold closure for [ʔ] could help resume phonation after a period of creak if the vocal folds are vibrating irregularly. There was a significant interaction between preceding glottalization and vowel-vowel hiatus. A hiatus environment (i.e. a vowel-initial word that was preceded by a word ending in a vowel) was found to be a significant predictor of full [ʔ] only when the preceding word ended in glottalization (i.e., with some form of irregularity).

Table 3: Results of logistic regression model predicting occurrence of [ʔ] vs. no glottalization for vowel-initial words.

	Coef β	SE(β)	**z**	**p**
Intercept	−6.15	1.24	−5.0	**<.0001**
Preceding break	1.19	0.14	8.4	**<.0001**
Hiatus=Y	−0.06	0.41	−0.1	>0.9
Accent=Y	0.41	0.31	1.3	>0.2
Prominence=Y	4.03	0.38	10.6	**<.0001**
Preceding glottalization=Y	1.26	0.32	3.9	**<.0001**
Following break	0.24	0.13	1.8	>0.1
Word type=function	−0.03	0.60	0.0	>1
Vowel frontness=front	0.26	0.64	0.4	>0.7
Vowel frontness=central	1.00	0.76	1.3	>0.2
Vowel height=low	0.57	0.52	1.1	>0.3
Vowel height=mid	−0.08	0.56	−0.1	>0.9
Vowel length=lax	0.06	0.54	0.1	>0.9
Preceding pause=Y	2.12	0.40	5.3	**<.0001**
Log freq. word	−0.22	0.19	−1.1	>0.3
Log freq. preceding word	−0.30	0.16	−1.9	>0.1
Preceding word type=function	1.15	0.47	2.4	**<.05**
Hiatus=Y:Preceding glottalization=Y	1.96	0.75	2.6	**<.01**

The final predictor to emerge as significant from the model (though much less so than prominence or preceding break) was the preceding word type. A preceding function word (compared to a content word) increased the likelihood of a word-initial vowel's being preceded by a full glottal stop, possibly because a glottal stop will render the vowel-initial word more prominent by preventing the function word from becoming a proclitic on the target word. For example, the sequence 'the only' is likely to be pronounced without a clear boundary between the determiner and the adjective ([ðioʊnli]). If produced with a full glottal stop ([ðəʔoʊnli]), the boundary between the determiner and the adjective is clearly defined, which might increase the degree of perceived prominence of the content word to listeners because the function word has not become a proclitic.

The relative importance of each of the significant factors was also assessed by comparing the full model to smaller models, each lacking one of the significant factors. This form of model comparison, done by means of the *anova* function in *R*, provides a chi-squared statistic and p-value indicating whether the full model provides a significantly better fit to the data than the model with a factor removed (Baayen 2008). The results mirror the z-scores of the estimates in the full model, indicating that the most important factors are, in order, prominence > preceding break > preceding glottalization > preceding pause > hiatus > preceding word type.

In sum, full glottal stops are more likely to occur when the vowel-initial word is phrase-initial

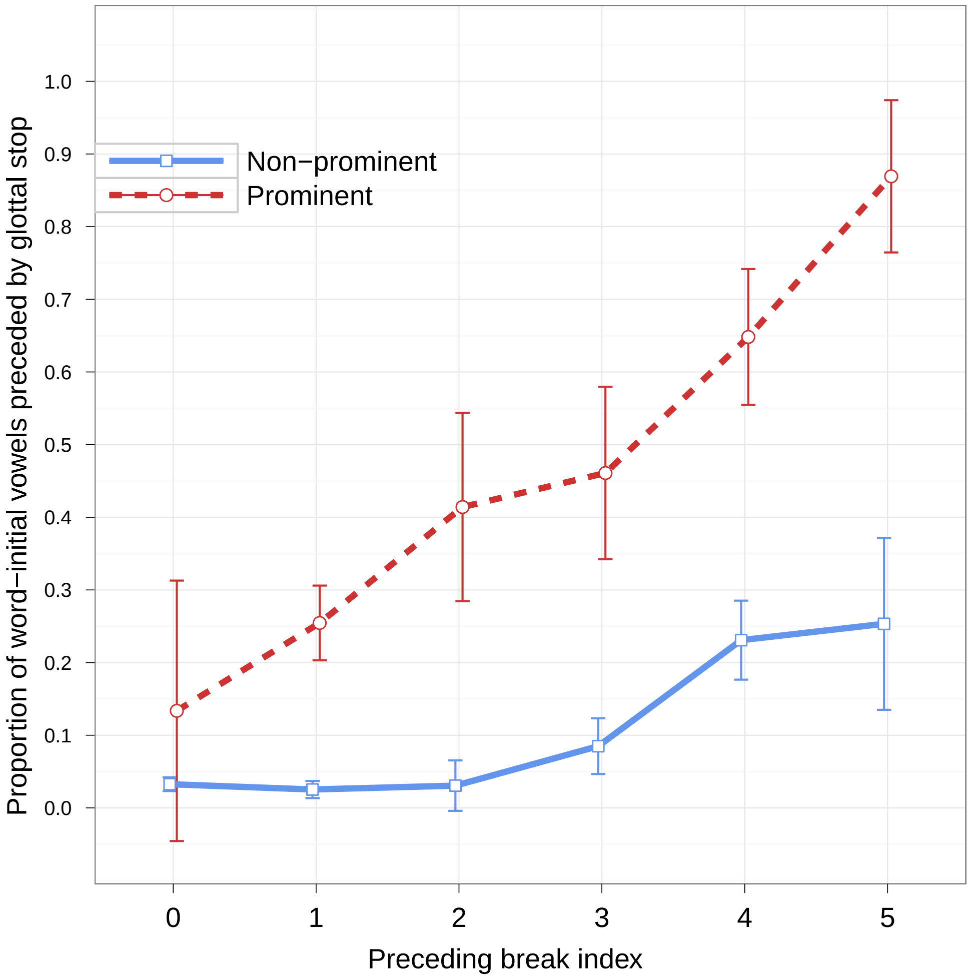

Figure 1: Proportion [ʔ] for vowel-initial words as a function of preceding break index and prominence. Error bars indicate 95% confidence intervals.

and when the vowel is prominent, which supports the findings by Dilley et al. (1996). Preceding pauses or glottalization, hiatus, and the preceding word type were also found to be significant predictors of full glottal stops, but much less so than prominence and preceding break index. By considering only the effects of prominence and phrasal position, it is possible to account for 95% of cases of [ʔ]. Prominence alone is able to account for three quarters of cases, and phrase-initial position for nearly seven of ten cases (Garellek 2012).

Knowing now that the vast majority of full glottal stops occur in prominent and/or phrase-initial environments, I turn to the cases of word-initial vowels *without* full glottal stops to determine if they show voice quality that is characteristic of laryngealization, i.e. voicing with more vocal fold closure. If they do, I will infer that such laryngealization in prominent and/or phase-initial environments is due to incomplete glottal stops ([ʔ̰]), because these same environments are known to be the most important factors in predicting full glottal stops.

2.2.2 Acoustic effects of prominence vs. phrasal strengthening on word-initial vowels

If prosody and phrasing are the most important predictors of full [ʔ], I hypothesize that they should also be good predictors of incomplete [ʔ̰] as well. To test this, I look at the voice quality of vowels without a full [ʔ], to see if they show characteristics of laryngealization, which would be consistent with the presence of an incomplete glottal stop. Note that by 'laryngealization' I refer specifically to voice quality with increased glottal closure (Gordon & Ladefoged 2001, Esling, Fraser & Harris 2005).

Recall from Section 2.1.2 that H1*-H2* is taken to be the likeliest acoustic measure to be correlated with increased glottal closure. Indeed, lower values of H1*-H2* are found for vowels following full [ʔ] (Garellek 2012). To test if lower values of H1*-H2* are associated with prominence on vowel-initial words that are not preceded by [ʔ], I fitted a linear mixed-effects model predicting H1*-H2* as a function of the prominence and phrasing. H1*-H2* (standardized and outliers with an absolute Z-score > 3 removed) is the dependent variable, and phrasal condition*prominence (main effects and interaction) are the fixed effects. Vowels preceded by a break index (BI) of 0 or 1 were coded as 'ip-medial,' those preceded by BI 3 and 4 were (respectively) 'ip-initial' and 'IP-initial,' and those preceded by BI 5 were 'Utterance-initial.' I excluded vowels preceded by a break index of 2, because these are unclear cases that could be phrase-medial or phrase-initial. Random intercepts are included for speaker and word, as well as a random slope of mean F0 by speaker. The prosodic condition factor is first coded using forward difference coding, such that (keeping prominence constant), the mean of each prosodic condition is compared to the mean of the following (higher) condition. Additionally, the difference in prominence across all prosodic conditions ('main effect' of prominence) is assessed separately. Pairwise comparisons between prominent vs. non-prominent sounds at a given prosodic domain are assessed by changing the contrast coding of the prosodic condition to zero-sum coding and subsequent reference changes. Significance of a pairwise difference in means is evaluated based on whether the absolute t-value was greater than 2, given that MCMC sampling has not yet been implemented for models with random slopes (Baayen 2008). Phrasal condition is also recoded as a linear factor (by increasing break index), in order to assess its main effect within prominence category. The results for word-initial vowels are shown in Figure 2. H1*-H2* is plotted as a function of prominence and phrasal domain. No difference as a function of prosodic domain is found. The difference in H1*-H2* as a function of vowel prominence is significant for ip-medial and IP-initial vowels. The main effect of prominence on H1*-H2* (across all phrasal domains) is also significant. T-values can be found in Table 4.

Table 4: T-values for pairwise comparisons of (standardized) H1*-H2* within and across prominence groups for word-initial vowels not preceded by [ʔ] (see also Figure 2). Values below -2 or greater than 2 are considered significant.

(a) Difference in H1*-H2* within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	-0.62	0.77	-0.12	0.12
Non-prominent	0.29	-1.51	-0.29	-1.61

(b) Difference in H1*-H2* for prominent vs. non-prominent initial vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
-4.04*	-1.04	**-2.80***	-1.08	**-2.78***

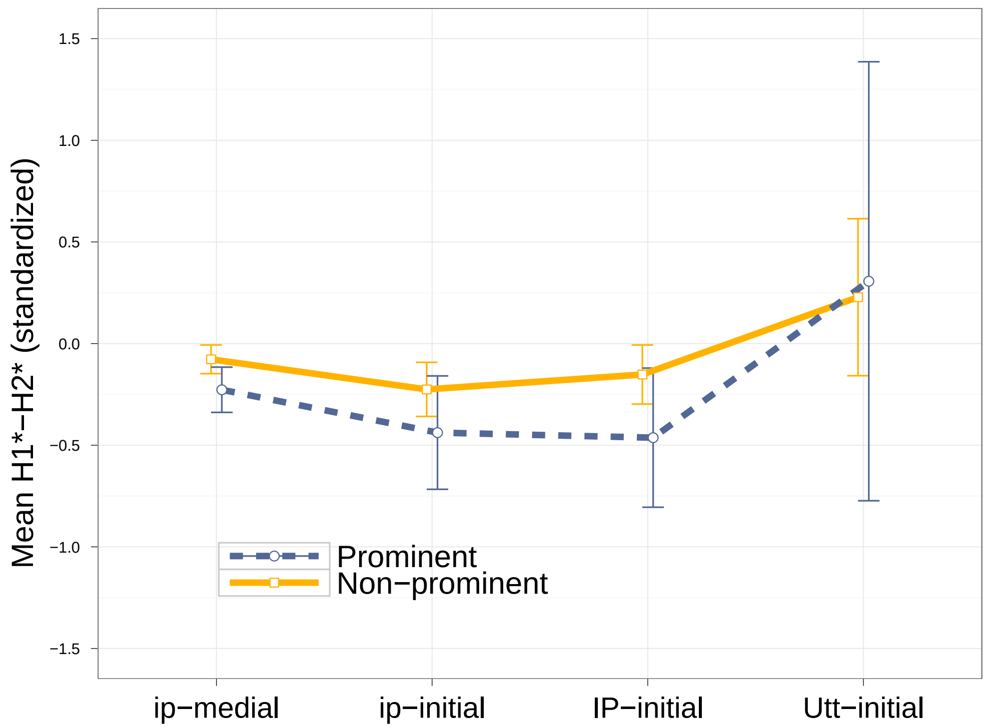

Figure 2: Mean H1*-H2* for word-initial vowels. Prominent vowels are the dashed lines. Error bars indicate 95% confidence intervals.

These results are consistent with the idea that prominent word-initial vowels – even those with no [ʔ] – are produced with increased vocal fold adduction. Surprisingly, higher prosodic domains are not associated with a decrease in H1*-H2*. This is inconsistent with the assumption that higher prosodic phrases trigger an increase in glottal stops. Therefore, for word-initial vowels that are not preceded by [ʔ], prominence induces greater laryngealization (based on lower values of H1*-H2*), whereas higher prosodic domains do not.

If prominence (rather than prosodic domain) is associated with incomplete glottal stops, why are full glottal stops [ʔ] more likely phrase-initially? Based on the acoustic findings of this section, it is likely that prominence is responsible for the presence of a glottal stop gesture, which may be realized either incompletely as [ʔ̰] or as full glottal stop [ʔ]. On the other hand, phrasal position accounts mostly for initial strengthening, which results in more instances of *full* glottal stops phrase- and utterance-initially than phrase-medially.

2.2.3 Acoustic effects of prominence vs. phrasal strengthening on word-initial sonorants and their following vowels

If the lowering of H1*-H2* for prominent vowel-initial words is due to a glottal stop gesture, then we would expect that such lowering would not occur for word-initial sonorants or their following vowels, because these positions are never preceded by a glottal stop in English. To test this, I fitted linear mixed-effects models to word-initial sonorants and then to their following vowels, in order to predict H1*-H2* or H1-H2 as a function of the prominence and phrasing. This was done using models identical in structure to those fitted to the word-initial vowels. However, for word-initial sonorants, H1-H2 (uncorrected for vowel formants) was used as the dependent variable, because formant tracking errors during the sonorants affected the values of (corrected) H1*-H2* (Garellek 2012).

The results for word-initial sonorants are shown in Figure 3a and Table 5). No significant main effect of prominence is found, but the main effect of phrasing is significant for both prominent and non-prominent sonorants. Higher phrasal domains are generally associated with an increase in H1-H2, even if pairwise differences between domains are not always significant. The results for word-initial sonorants are shown in Figure 3b. No significant main effects or interactions are found (see Table 6).

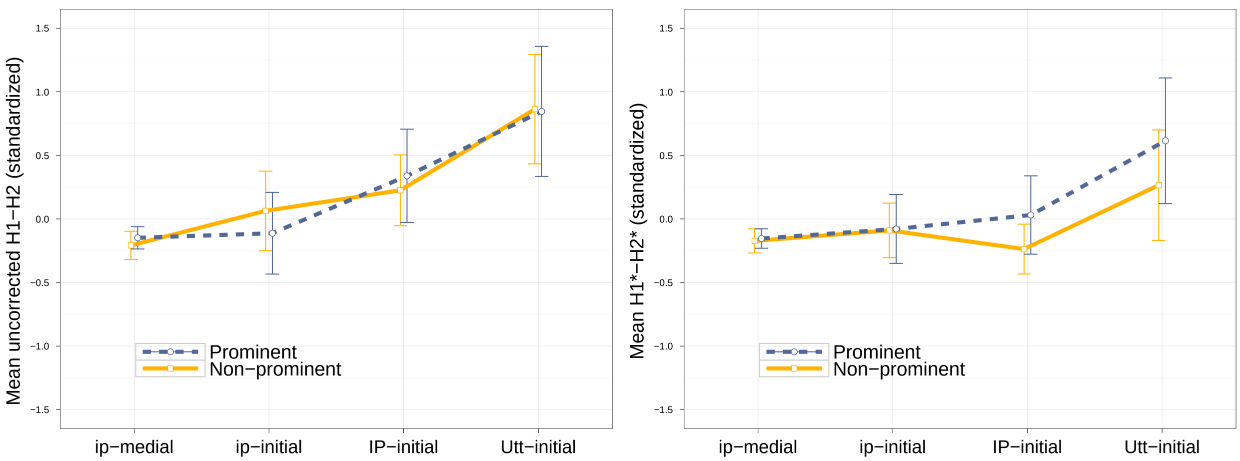

Figure 3: Mean H1*-H2* (or H1-H2) for word-initial sonorants (left), and post-sonorant vowels (right). H1-H2 is used for initial sonorants because of problems with formant tracking that affect corrected measures (Garellek 2012). Prominent vowels are the dashed lines. Error bars indicate 95% confidence intervals.

Table 5: T-values for pairwise comparisons of (standardized) uncorrected H1-H2 within and across prominence groups for word-initial sonorants (see also Figure 3a). Values below -2 or greater than 2 are considered significant.

(a) Difference in H1-H2 within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	-0.17	**2.62***	1.23	**5.09***
Non-prominent	**3.35***	1.84	**2.76***	**3.59***

(b) Difference in H1*-H2* for prominent vs. non-prominent initial vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
-0.57	0.88	-0.34	-1.59	-1.91

These results show that, unlike for word-initial vowels, H1*-H2*/H1-H2 is not lower for prominent initial sonorants or post-sonorant vowels. This is consistent with the hypothesis that only word-initial vowels should show laryngealization, because only they are preceded by a glottal stop

Table 6: *T*-values for pairwise comparisons of (standardized) H1*-H2* within and across prominence groups for post-sonorant vowels (see also Figure 3b). Values below -2 or greater than 2 are considered significant.

(a) Difference in H1-H2 within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	0.03	-0.58	-0.66	0.51
Non-prominent	-1.34	0.57	-0.95	1.50

(b) Difference in H1*-H2* for prominent vs. non-prominent initial vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
0.07	1.04	-0.05	0.10	-0.57

gesture. Crucially, the increase in H1-H2 as a function of phrasing for word-initial sonorants suggests that phrase-initial voicing is generally breathier, not more laryngealized.

2.3 Discussion of corpus study

This study seeks to answer two questions regarding English word-initial glottal stops:

1. Which factors matter most in predicting the occurrence of glottal stops?
2. Is word-initial glottalization due to voice quality strengthening?

The results from Section 2.2.1 show that [ʔ] is predicted largely by prominence and phrase-initial position, which together can account for 95% of cases of full glottal stops. This finding is in line with previous researchers, who have shown that prominence (accent or stress) and phrasing are important in predicting glottal stops and/or glottalization (Pierrehumbert & Talkin 1992, Pierrehumbert 1995, Dilley et al. 1996, Redi & Shattuck-Hufnagel 2001, Davidson & Erker 2012).

To find cases of glottal stops that have been realized incompletely, in Section 2.2.2 I looked at the effects of prominence and phrasal position on the acoustic properties of word-initial vowels. I focused on these two factors because they are by far the most important factors in predicting full [ʔ] occurrence, and are therefore also likely to predict occurrences of incomplete [ʔ̞]. The results show that prominence can in fact be used to predict laryngealization that is typical of voicing with increased glottal closure. However, this is only true for word-initial vowels. This result is expected if prominent word-initial vowels are likely to be preceded by a glottal stop gesture. In contrast, word-initial sonorants and their following vowels do not show laryngealized phonation when prominent. Indeed, they are also never preceded by a full glottal stop.

On the other hand, phrasing does not seem to affect the voice quality of initial vowels differently than that of initial sonorants. Instead, all voiced segments show noisier phonation at the onset of higher prosodic domains. This does not mean that phrasing never accounts for the presence of a glottal stop; about 20% of phrase-initial but non-prominent vowels are preceded by full glottal stops, and about the same percentage of cases of [ʔ] can be explained by phrasing and not by prominence. However, in general, onsets of higher prosodic domains yield noisier phonation that is more characteristic of breathiness or of creak with increased vocal fold abduction (cf. Slifka (2006)).

3 Articulatory study

The main goals of this laboratory speech study are to confirm (by means of electroglottograhic contact quotient) the interpretation of the acoustic findings from the corpus study, and to see if they can be generalized to another language (here, Spanish). In Spanish, word-initial glottalization is known to be rarer than in English (Bissiri et al. 2011), but the phrasing effects found in the previous study might generalize to Spanish. Thus, the specific hypotheses tested are:

1. In English, word-initial vowels should show increased contact under prominence. Word-initial sonorants or post-sonorant vowels should not.

2. In Spanish, prominence should result in a less contact for vowel-initial words than in English, because glottalization is rarer in Spanish.

3. In both languages, higher phrasal domains should show decreased contact.

I will also obtain simultaneous acoustic measures (of H1*-H2*), which will allow for replication of the corpus study in laboratory speech. If the acoustic results mirror those from the corpus study, then I assume that the articulatory measures here can also be used to interpret the corpus results.

3.1 Method

3.1.1 Stimuli

The target words in both languages consisted of vowel- or sonorant-initial proper nouns with two or three syllables. Each target sound appeared with and without primary stress. Target word-initial vowels were [æ, ə, oʊ], and target word-initial sonorants were [m, n, l, j, w]. Sonorant-initial words were followed by the vowels [oʊ], [ɔ], or [ɪ] in English and by [a], [o], or [au] in Spanish. Stressed syllables were intended to attract phrasal prominence by means of a pitch accent, whereas unstressed syllables were considered non-prominent. In total, there were 16 English target names and 14 Spanish names. The complete list of target words in both English and Spanish is shown in Table 7.

Each target word was placed in four distinct positions that were likely to be pronounced in four distinct phrasal positions: Utterance-initially (after a breath), IP-initially after a high boundary tone (H%), ip-initially after a high phrase accent (H-), and ip-medially. In the three Utterance-medial conditions, the target words always followed a fixed vowel ([ɚ] in English and [a] in Spanish). Additionally, in both languages the number of syllables preceding the target word (when utterance-medial) was held constant, as was the total number of syllables (per stress condition), with the exception of the trisyllabic English names *Yolanda* and *Winona*. Thus, in both stress conditions the target syllable was the seventh syllable in the utterance if it occurred utterance-medially. Utterances with stressed target sounds had a total of 15 syllables, whereas those with unstressed target sounds had 15 or 16 in English and 16 in Spanish. The target syllable, if stressed, never bore the nuclear pitch accent of the phrase. The sentence frames in both English and Spanish are shown in Table 8. The expected breaks and preceding tones in MAE-ToBI (Beckman & Ayers Elam 1997) and Mexican Spanish ToBI (de-la-Mota, Butragueño & Prieto 2010) are indicated.

3.1.2 Participants

In total, 24 participants were recruited: 12 native speakers of American English (six female and six male), and 12 native speakers of Mexican Spanish (seven female and five male). Both English

Table 7: Target words in English and Spanish. The target sounds are underlined.

(a) Target words in English.

	Sound	Stressed syllable	Unstressed syllable
Initial vowel	[æ, ə]	*<u>A</u>nna*	*<u>A</u>nnette*
	[i]	*<u>I</u>gor*	*<u>Y</u>vette*
	[oʊ]	*<u>O</u>din*	*<u>O</u>dette*
Initial sonorant	[m]	*<u>M</u>organ*	*<u>M</u>aureen*
	[n]	*<u>N</u>ora*	*<u>N</u>oreen*
	[l]	*<u>L</u>aura*	*<u>L</u>oraine*
	[j]	*<u>Y</u>oko*	*<u>Y</u>olanda*
	[w]	*<u>W</u>innie*	*<u>W</u>inona*

(b) Target words in Spanish.

	Sound	Stressed syllable	Unstressed syllable
Initial vowel	[a]	*<u>A</u>na*	*<u>A</u>nita*
	[e]	*<u>E</u>va*	*<u>E</u>vita*
	[o]	*<u>O</u>lga*	*<u>O</u>limpia*
Initial sonorant	[m]	*<u>M</u>aya*	*<u>M</u>arina*
	[n]	*<u>N</u>ana*	*<u>N</u>anita*
	[l]	*<u>L</u>aura*	*<u>L</u>aurita*
	[j, ɟ]	*<u>Y</u>ola*	*<u>Y</u>olanda*

and Spanish speakers were UCLA students, and were awarded course credit for their participation. All native English speakers spoke only English fluently. The average participant's age was 21 (SD = 2.6) for English speakers and 22 (SD = 2.3) for the Spanish speakers. The native Spanish speakers also spoke English, though their levels of proficiency in English varied. Five Spanish speakers were raised in Mexico; the remaining speakers were born in the Los Angeles area. All spoke Spanish on a daily basis, and all claimed to be equally or more comfortable speaking Spanish compared to English. The Spanish-speaking participants spoke the *Distrito Federal* variety of Mexican Spanish. To ensure that none of the Spanish speakers spoke English-accented Spanish, the Spanish recordings were labeled by a native speaker who confirmed that all speakers spoke unaccented Mexican Spanish.

3.1.3 Procedure

The task consisted of recorded read speech in either English or Spanish. Each participant read all the target words in four phrasal conditions, and each sentence was repeated twice. Thus, each native English speaker said 60 sentences twice for a total of 120 repetitions, and each native Spanish speaker read 56 sentences twice, for a total of 112 repetitions. The order of the sentences was randomized, such that no two participants read the sentences in the same order. Participants were instructed to read each sentence aloud as naturally as possible, with no special emphasis on a particular word. The recordings were not ToBI-labeled, but labelers listened to them and noted if the preceding break or tone differed from the intended category (as shown in Table 8) in any repetition, in which case that token was excluded. If participants read a sentence with focus on the target word, such that it bore a nuclear pitch accent, or if they read a sentence with too small

Table 8: Sentence frames in English and Spanish. The location of the target word is marked by '*X*'. See Table 7 for target words.

(a) Sentence frames in English.

Phrasal condition	Preceding tone	Preceding break index	Sentence frame
Utterance-initial	(none: breath)	'5'	X *was sitting on the sofa for the entire day.*
IP-initial	H%	4	*Was that Alexander?* X *was talking to him today.*
ip-initial	H-	3	*Teddy, Alexander,* X*'s older sister, and Jim slept.*
ip-medial	(L+)H*	1	*Alex liked to bother* X*'s older sister on the trip.*

(b) Sentence frames in Spanish.

Phrasal condition	Preceding tone	Preceding break index	Sentence frame and translation
Utterance-initial	(none: breath)	'5'	X *estuvo sentada casi todo el día.* 'X was sitting for nearly the whole day.'
IP-initial	H%	4	*¿Viste a María?* X *no puede encontrarla.* 'Did you see María? X can't find her.'
ip-initial	H-	3	*Paulina Rivera,* X *Cuarón y yo nos fuimos.* 'Paulina Rivera, X Cuarón, and I left.'
ip-medial	L+(>)H*	1	*En el parque vi a* X *con su hermano Pedro.* 'I saw X in the park with her brother Pedro.'

or large a preceding break, they were asked to repeat the sentence. Any readings with unexpected Utterance-medial breaks or dysfluencies were also excluded, and participants were then asked to repeat the sentence.

The participants were recorded in a sound-attenuated room at UCLA using a Shure SM10A head-mounted microphone in Audacity at a sampling rate of 22,050 Hz. Simultaneous EGG recordings (as a second channel in a stereo audio recording) were collected at the same sampling rate using a two-channel Glottal Enterprises electroglottograph (Model EG2), with a high-pass filter of 20 Hz. The recording lasted approximately 30-45 minutes.

3.1.4 Labeling and measures

The target sounds were labeled and extracted for acoustic and electroglottographic analyses. In the case of word-initial sonorants, the post-sonorant vowel was also extracted for subsequent analysis, following the same criteria used in the corpus study (and extended to Spanish). For ip-medial sentences with *Ana* in Spanish (in the sentence *En el parque vi a Ana...*), often the first [a] (from the preposition 'a') was difficult to distinguish from the [a] in 'Ana' (indeed, the preposition was often elided). If the labelers heard the two instances of [a] as distinct, the boundary was taken to be the middle of the long [a] sequence. Otherwise, if they head only one [a], it was attributed to the target word.

The audio waveforms of the extracted segments were then analyzed for acoustic measures using VoiceSauce (Shue et al. 2011), as described above. The electroglottographic waveforms were analyzed for EGG measures using EggWorks, a free EGG analysis program created by Henry Tehrani at UCLA. The subsequent analysis will focus on electroglottographic contact quotient as the articulatory correlate of vocal fold contact, and H1*-H2* (or uncorrected H1-H2 in the case of sonorants) as the acoustic correlate of vocal fold contact. Contact quotient (CQ) was measured using the hybrid method, which defines the point of vocal fold closure as the peak in the derivative of the EGG signal (following Howard (1995)), and uses a 25% (peak-to-peak amplitude) threshold for detecting the point of vocal fold opening (following Orlikoff (1991)). This hybrid method for measuring contact was used because thresholds at 20% and 25% are found to be best correlated with contact measured via direct imaging of the glottis (Herbst & Ternström 2006), and because this particular version of EGG contact quotient was found to be most sensitive to changes in voice quality (Kuang 2011). CQ, H1*-H2*, and H1-H2 were standardized within speaker, and outliers (absolute Z-score > 3) were removed.

3.2 Results

3.2.1 English results

In the analyses that follow, the statistical difference in mean value for H1*-H2* or H1-H2 is assessed by linear mixed-effects modeling. As in the corpus study, the acoustic or EGG measure is the dependent variable, and prosodic condition*prominence (main effects and interaction) are the fixed effects. The models also had the same random structure as the models in Section 2). The results for word-initial vowels are shown in Figure 4. CQ (top) and H1*-H2* (bottom) are plotted as a function of prominence and phrasal domain. For prominent vowels, no difference as a function of prosodic domain is found. For non-prominent vowels, there is a significant main effect of phrasing, whereby higher domains are associated with lower values of CQ. The difference in CQ as a function of vowel prominence is significant at all phrasal levels above the ip. The main effect of prominence on CQ (across all phrasal domains) is also significant. Surprisingly, none of the differences in H1*-H2* between prominent vs. non-prominent vowels was significant, as can be seen in Figure 4b (cf. results from Chapter 2). T-values can be found in Table 9.

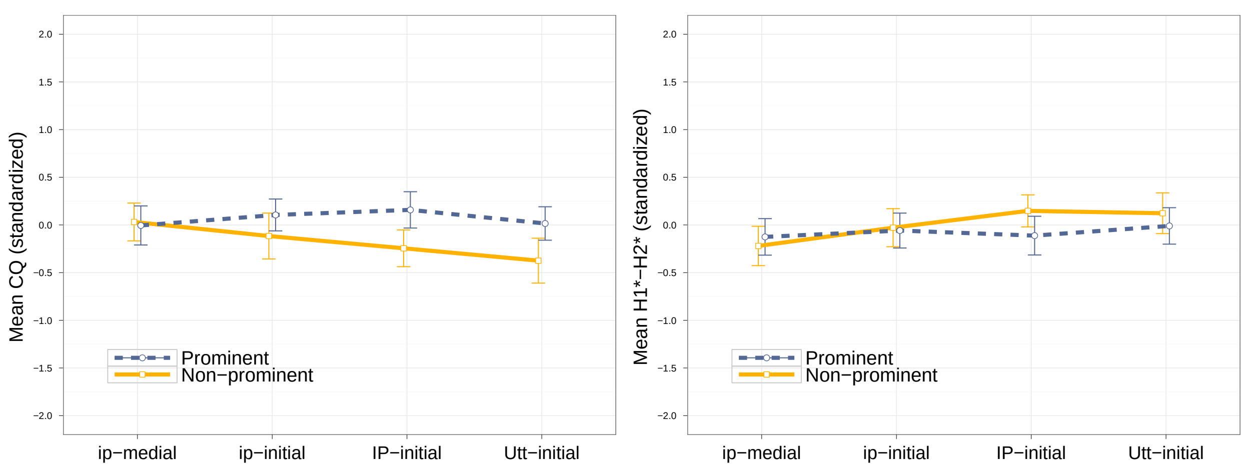

Figure 4: Mean CQ (left) and H1*-H2* (right) for word-initial vowels in English. Prominent vowels are the dashed lines. Error bars indicate 95% confidence intervals.

The results for word-initial sonorants are shown in Figure 5. Recall that uncorrected H1-H2 is used here because of problems with the formant correction during sonorants. Regardless of prominence, ip-medial sonorants have higher contact and lower values of H1-H2 than ip-initial

Table 9: T-values for pairwise comparisons of (standardized) CQ and H1*-H2* within and across prominence groups for word-initial vowels in English (see also Figure 4). Values below -2 or greater than 2 are considered significant.

(a) Difference in CQ within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	-0.50	-0.50	0.94	-0.22
Non-prominent	0.83	1.02	1.06	**-2.80***

(b) Difference in CQ for prominent vs. non-prominent initial vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
0.22	1.56	**3.21***	**2.98***	**3.97***

(c) Difference in H1*-H2* within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	-0.34	1.04	-0.32	0.42
Non-prominent	-1.22	-0.54	-0.77	1.22

(d) Difference in H1*-H2* for prominent vs. non-prominent initial vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
0.14	-0.12	-0.57	-0.72	-0.33

sonorants (see Table 10). In addition, ip-initial sonorants have higher CQ values than IP-initial ones, though no corresponding change in H1-H2 is found. The main effects of phrasing on CQ and H1-H2 are significant for prominent and (for H1-H2) for non-prominent sonorants. Higher prosodic domains are associated with a decrease in CQ and an increase in H1-H2. No effect of prominence is found on either CQ or H1-H2.

We now turn to the vowels following initial sonorants. Recall that these vowels are meant to provide two kinds of information: (1) whether phrasal effects last throughout the initial syllable, and (2) whether non-initial vowels behave similarly to initial ones, with regards to both prominence and phrasing effects. In terms of phrasing effects, ip-medial post-sonorant vowels have higher CQ contact than ip-initial ones. There is also a main effect of phrasing on CQ for prominent post-sonorant vowels, with a decrease in CQ as a function of increasingly-high phrasal domain. Similarly to what was found in the corpus study, no phrasing effects on H1*-H2* are significant. No effects of prominence on CQ are significant, but significantly lower H1*-H2* values under prominence are found in Utterance-initial position. However, no main effect of prominence is found for either measure. The results support the assumption that phrasal effects will last beyond the word-initial segment; there is a significant drop in CQ as a function of an increasingly-high prosodic domain (but only for prominent vowels). That post-sonorant vowels do not show a main effect of prominence is similar to the result in the corpus study. However, effects of prominence on H1*-H2* are indeed found, though only at the highest phrasal position.

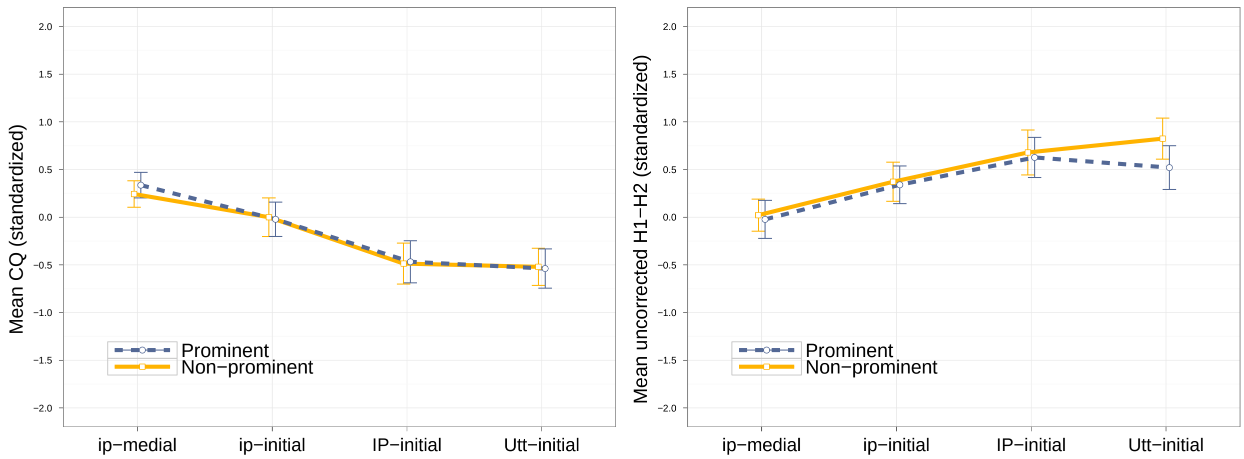

Figure 5: Mean CQ (left) and H1-H2 (right) for word-initial sonorants in English. Prominent sonorants are the dashed lines. Error bars indicate 95% confidence intervals.

Table 10: T-values for pairwise comparisons of (standardized) CQ and H1-H2 within and across prominence groups for word-initial sonorants in English (see also Figure 5). Values below -2 or greater than 2 are considered significant.

(a) Difference in CQ within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	**2.94***	**4.10***	0.51	**-2.62***
Non-prominent	**2.05***	**3.53***	0.45	-1.88

(b) Difference in CQ for prominent vs. non-prominent sonorants.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
0.72	0.09	-0.39	-0.39	0.10

(c) Difference in (uncorrected) H1-H2 within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	**-2.32***	-0.08	0.91	**2.85***
Non-prominent	**-2.89***	-1.28	0.23	**2.89***

(d) Difference in (uncorrected) H1-H2 for prominent vs. non-prominent sonorants.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
0.24	-0.11	-0.90	-1.30	-0.64

3.2.2 Spanish results

The results for word-initial vowels in Spanish are shown in Figure 7. CQ (left) and H1*-H2* (right) are plotted as a function of prominence and phrasal domain. For both prominent and non-prominent vowels, higher values of CQ and lower values of H1*-H2* are found ip-medially

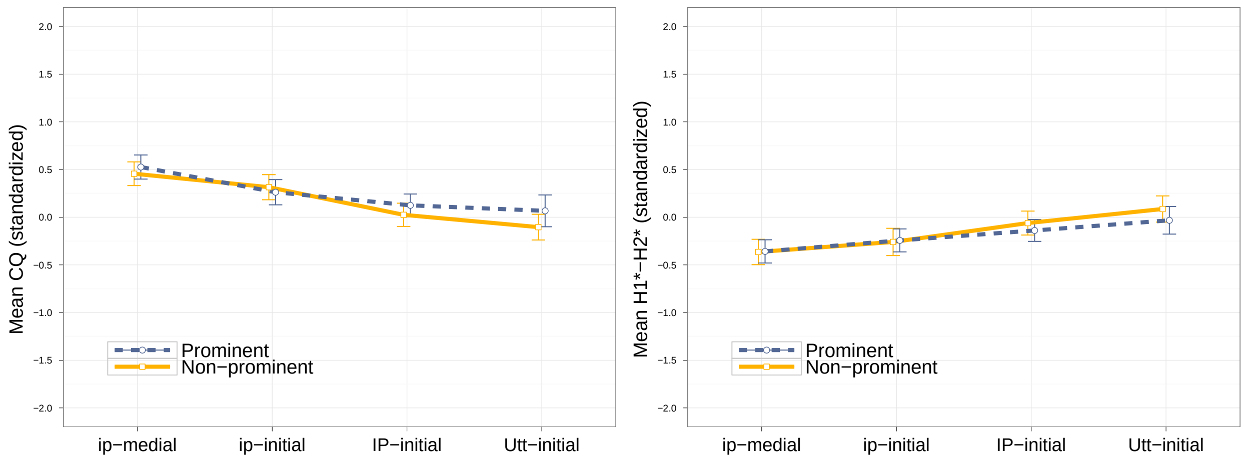

Figure 6: Mean CQ (left) and H1*-H2* (right) for post-sonorant vowels in English. Prominent vowels are the dashed lines. Error bars indicate 95% confidence intervals.

Table 11: *T*-values for pairwise comparisons of (standardized) CQ and H1*-H2* within and across prominence groups for post-sonorant vowels in English (see also Figure 6). Values below -2 or greater than 2 are considered significant.

(a) Difference in CQ within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	**3.04***	0.22	0.03	**-2.86***
Non-prominent	**2.00***	**2.14***	0.87	-1.72

(b) Difference in CQ for prominent vs. non-prominent post-sonorant vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
0.55	-0.07	1.11	1.60	0.93

(c) Difference in H1*-H2* within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	-0.82	1.69	0.58	1.00
Non-prominent	-1.12	0.02	-0.30	0.99

(d) Difference in H1*-H2* for prominent vs. non-prominent post-sonorant vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
-0.37	-0.58	-1.83	**-2.44***	-1.67

compared to ip-initial position, and there is a significant corresponding main effect of phrasing on both measures. CQ is higher for non-prominent Utterance-initial vowels than for non-prominent IP-initial vowels, likely indicating increased glottalization of non-prominent initial vowels at the highest prosodic domain. Prominent initial vowels have increased CQ and lower H1*-H2* IP- and Utterance-initially. Prominent vowels have lower values of H1*-H2* also at the ip-initial level.

There is an overall main effect of prominence on CQ, with higher values found for prominent vowels than for non-prominent ones. A corresponding main effect of prominence is found for H1*-H2*, with lower values of the measure under prominence. T-values can be found in Table 12.

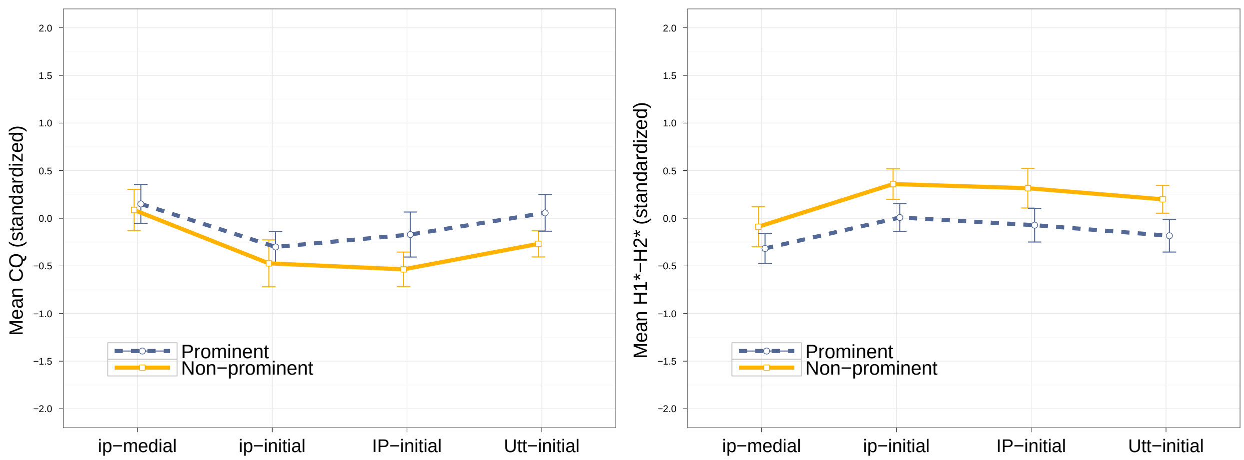

Figure 7: Mean CQ (left) and H1*-H2* (right) for word-initial vowels in Spanish. Prominent vowels are the dashed lines. Error bars indicate 95% confidence intervals.

Table 12: T-values for pairwise comparisons of (standardized) CQ and H1*-H2* within and across prominence groups for word-initial vowels in Spanish (see also Figure 7). Values below -2 or greater than 2 are considered significant.

(a) Difference in CQ within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	**3.44***	-1.02	-1.63	**-3.79***
Non-prominent	**4.61***	0.12	**-2.06**	**-5.03***

(b) Difference in CQ for prominent vs. non-prominent initial vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
0.37	1.55	**2.53***	**2.23***	**2.61***

(c) Difference in H1*-H2* within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	**-2.56***	1.42	0.91	**2.74***
Non-prominent	**-4.52***	1.43	0.00	**4.59***

(d) Difference in H1*-H2* for prominent vs. non-prominent initial vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
-0.53	**-2.65***	**-2.77***	**-2.65***	**-2.13***

The results for word-initial sonorants in Spanish are shown in Figure 8 and Table 13. CQ (left) and (uncorrected) H1-H2 (right) are plotted as a function of prominence and phrasal domain. When non-prominent, ip-medial sonorants have higher values of CQ than ip-initial ones, though both prominent and non-prominent ip-initial sonorants have corresponding lower values of H1-H2 than ip-initial ones, which in turn have lower values than IP-initial ones. Non-prominent sonorants also show a significant decrease in H1-H2 at the Utterance-initial level. A significant main effect of phrasing on CQ is found for non-prominent sonorants, with CQ decreasing as a function of higher prosodic domain. The main effect of phrasing on H1-H2 is significant for both prominent and non-prominent sonorants.

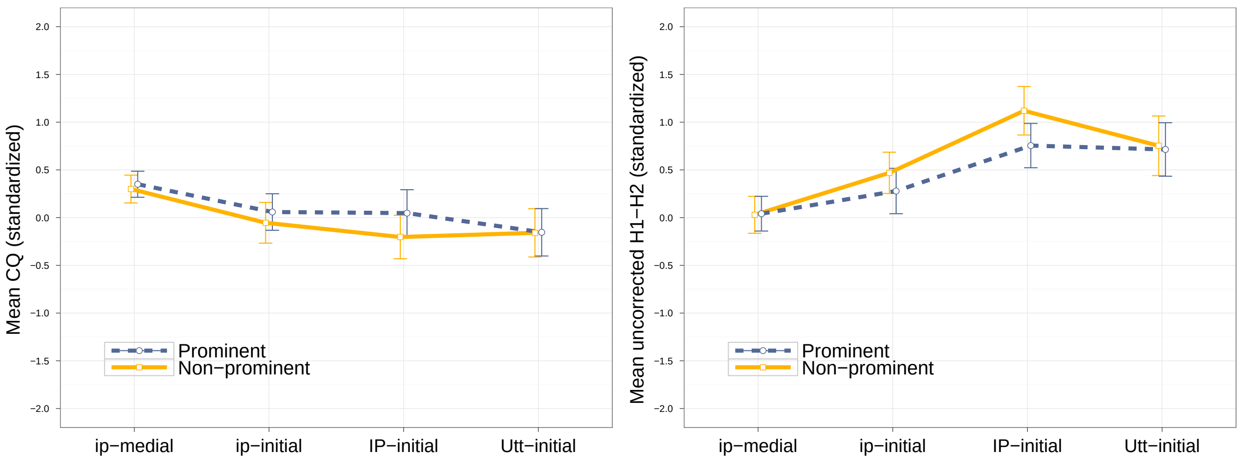

Figure 8: Mean CQ (left) and H1-H2 (right) for word-initial sonorants in Spanish. Prominent sonorants are the dashed lines. Error bars indicate 95% confidence intervals.

The results for post-sonorant vowels in Spanish are shown in Figure 9. CQ (left) and H1*-H2* (right) are plotted as a function of prominence and phrasal domain. The only significant effect of phrasing is the decrease in CQ from prominent ip-medial post-sonorant vowels to prominent ip-initial vowels. No significant main effect of phrasing is found. At no prosodic level is there a significant effect of prominence on CQ, and no main effect of prominence on CQ is found. However, the results for H1*-H2* reveal that ip-medial and IP-initial post-sonorant vowels show lower values of the measure under prominence. There is also a significant main effect of prominence overall, with prominent post-sonorant vowels having lower values of H1*-H2* than non-prominent ones. T-values can be found in Table 14.

3.3 Discussion of articulatory study

Four hypotheses were outlined at the start of this study. First, word-initial vowels (but not initial sonorants) in English and Spanish should show increased contact under prominence. Second, prominence in Spanish should result in a smaller increase in contact for vowel-initial words than in English, because glottalization is rarer in Spanish. Lastly, in both languages, higher phrasal domains should show decreased contact. The results from this study support these hypotheses. Generally, word-initial vowels that are prominent show increased vocal fold contact, though the effect is stronger in English. Phrase-initial voicing in both English and Spanish is characterized by a decrease in contact. In the following section, I discuss the results from this study and the corpus study in Section 2, and their implications of both studies for theories of word-initial glottalization.

Table 13: *T*-values for pairwise comparisons of (standardized) CQ and H1-H2 within and across prominence groups for word-initial sonorants in Spanish (see also Figure 8). Values below -2 or greater than 2 are considered significant.

(a) Difference in CQ within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	1.55	0.39	1.52	-1.20
Non-prominent	**2.32***	1.43	-0.33	**-2.35***

(b) Difference in CQ for prominent vs. non-prominent sonorants.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
0.45	1.16	**2.07***	0.21	1.77

(c) Difference in (uncorrected) H1-H2 within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	**-2.69***	**-2.46***	1.68	**3.18***
Non-prominent	**-3.81***	**-3.62***	**2.55***	**4.32***

(d) Difference in (uncorrected) H1-H2 for prominent vs. non-prominent sonorants.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
-0.32	-0.88	-1.46	-0.83	-1.02

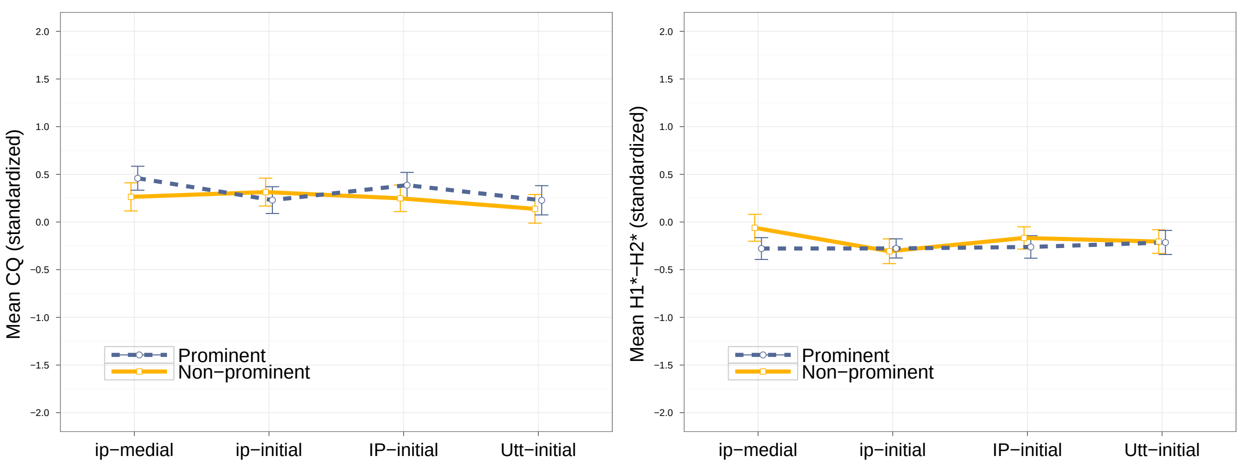

Figure 9: Mean CQ (left) and H1*-H2* (right) for post-sonorant vowels in Spanish. Prominent vowels are the dashed lines. Error bars indicate 95% confidence intervals.

Table 14: T-values for pairwise comparisons of (standardized) CQ and H1*-H2* within and across prominence groups for post-sonorant vowels in Spanish (see also Figure 9). Values below -2 or greater than 2 are considered significant.

(a) Difference in CQ within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	**2.54***	-1.90	1.44	-1.98
Non-prominent	-0.79	0.55	1.30	1.18

(b) Difference in CQ for prominent vs. non-prominent post-sonorant vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
1.43	-0.76	0.85	0.74	0.69

(c) Difference in H1*-H2* within prominence.

	ip-medial vs. ip-initial	ip-initial vs. IP-initial	IP-initial vs. Utt-initial	Main effect-phrasing
Prominent	-0.53	1.60	0.67	0.22
Non-prominent	1.77	-0.75	1.47	-1.44

(d) Difference in H1*-H2* for prominent vs. non-prominent post-sonorant vowels.

ip-medial	ip-initial	IP-initial	Utt-initial	Main effect of prominence
-2.15*	-0.71	**-2.02***	-1.71	**-2.02***

4 General discussion

4.1 Effects of prominence on voice quality

The acoustic effects of prominence on voice quality from the corpus study are largely confirmed in the EGG study. In English, only word-initial vowels show a consistent effect of prominence on voice quality: increased contact. Although the two languages behave similarly, the effect of prominence on word-initial vowel contact and H1*-H2* is greater in degree and more robust across the prosodic hierarchy in English than it is in Spanish, supporting the idea that word-initial glottalization in English is more common than in Spanish (Bissiri et al. 2011). Glottalization of vowel-initial words can be seen as an intentional form of prosodic (specifically, prominence) strengthening. To mark an initial vowel as prominent, speakers deliberately render it more consonantal (Pierrehumbert & Talkin 1992).

Alternatively, word-initial glottalization as a form of prominence strengthening could be an unintentional result of a more forceful articulation (Fougeron 2001; citing also Straka (1963) and Fujimura (1990)). Thus, word-initial glottal stops could be produced because prominent vowels are more forcefully articulated, not because of any linguistically-motivated purpose like contrast enhancement. However, this line of analysis is problematic, because we should also expect prominent sonorants and even non-prominent voiced segments to show increased contact. Thus, under this account it is difficult to explain the different effects found for vowels vs. sonorants and for prominence vs. initial strengthening.

4.2 Effects of phrasing on voice quality

The results of the two studies indicate that vowel and sonorant voicing at the onsets of higher prosodic domains shows a decrease in contact quotient and a corresponding increase in H1*-H2* or H1-H2. Why would voicing show less contact at higher phrasal domains? In this section, I propose several possibilities for why phrase-initial abduction occurs.

Initial strengthening is a possible explanation for phrase-initial abduction. However, strengthening that results in a more forceful articulation (Fougeron 2001) cannot be the origin of the phrasal effects found here, because a more forceful articulation is expected to result in laryngealization, rather than in breathy-like phonation.

On the other hand, during initial strengthening the *contrast* between the target sound and its neighbors could be strengthened ('syntagmatic enhancement'; see discussion by Fougeron & Keating (1997), Hsu & Jun (1998), Fougeron (2001), and Cho (2005), among others). Syntagmatic enhancement does not necessarily result in a more forceful or prototypical articulation. For example, the decrease in duration and amount of nasal flow for phrase-initial nasals in French (Fougeron 2001) can be viewed as a form of syntagmatic enhancement: phrase-initial nasals are more consonantal and thus less similar to the following vowel, even though they become less 'nasal.' Likewise, because voice quality at phrasal onsets shows a decrease in contact, word-initial vowels and sonorants can be viewed as becoming more consonantal (breathier, more [h]-like). This interpretation however runs into problems when we compare word-initial sonorants to post-sonorant vowels. In English and (to a lesser degree) in Spanish, post-sonorant vowels also show decreased vocal fold contact at higher prosodic domains (see Figures 6 and 9). If the purpose of decreasing vocal fold contact phrase-initially is to enhance the contrast between the initial segment and what follows, then we cannot explain why the entire initial syllable – not just the initial segment – shows the effect.

Another possible explanation for the decrease in contact phrase-initially is that it is due to respiratory constraints on voicing initiation. Indeed, Slifka (2000) provides evidence that glottal area increases at voicing initiation. This might be deliberate on the part of the speaker, because an increase in glottal opening would reduce phonation onset time by increasing transglottal airflow. But the increase in glottal area might also be unintentional, in that during inspiration the vocal folds open widely to allow air to pass to the lungs. Thus, possible respiratory reasons for decreased vocal fold contact are (1) that speakers increase glottal opening to increase transglottal flow and decrease phonation onset time, or (2) that the glottis is more open Utterance-initially due to the preceding intake of breath. However, neither explanation based on respiratory constraints is entirely satisfactory, because all phrasal onsets show a decrease in contact. Indeed, the most consistent effect of phrasing on voice quality is the difference between ip-medial (word-initial) and ip-initial domains, which are both Utterance-medial. Onsets of Utterance-medial phrases should not be influenced by respiration, at least not for the reasons mentioned above. No pauses occurred at intermediate phrase boundaries, and the ip-initial vowels and sonorants were preceded by a voiced segment, meaning that voicing did not have to 'restart' at ip onsets.

It is also possible that the Utterance-initial effects are generalized and phonologized down the prosodic hierarchy. This would then be similar to coda obstruent devoicing: the aerodynamic/laryngeal conditions favoring coda-devoicing are only found Utterance-finally (Westbury & Keating 1986), yet the process often generalizes to all word-final or even syllable-final positions in some languages (see Myers (2012) for a review of languages). Generalization of decreased contact for Utterance-initial voicing to all phrasal onsets could provide listeners with a consistent perceptual cue to phrase onsets, though it is currently unclear whether listeners attend to such changes in voice quality.

Unrelated to respiratory constraints, a final possible reason for phrase-initial abduction is pitch reset. Pitch reset is the change in the slope of F0 declination (Ladd 1984; 2008). The domain of pitch reset is often assumed to be at the level of the ip, though there is evidence for increased pitch reset at higher domain onsets (O'Shaughnessy & Allen 1983, Ladd 1988; 2008). Crucially, all phrase onsets are accompanied by pitch reset. I propose that pitch reset could trigger increased vocal fold abduction if it involves muscular relaxation of the thyroarytenoid (TA) and/or cricoarytenoid (CT). Indeed, there is evidence of TA (vocalis) and CT relaxation when F0 falls (Hirano, Ohala & Vennard 1969), which we would expect given increased CT and TA activation at higher pitch levels (Hirose 1997, Kreiman & Sidtis 2011)). Unexpectedly though, Hirano et al. (1969) also found vocalis and CT relaxation before a sharp rise in F0. CT and TA relaxation results in a decrease in vocal fold adduction, as Mendelsohn & Zhang (2011) and Zhang (2011) show. Thus, the rapid change in F0 triggered by pitch reset at phrase boundaries could be responsible for increased vocal fold abduction at all phrasal domains.

In sum, the most likely explanations for phrase-initial effects on voice quality are respiratory constraints on Utterance-Initial phonation (if then phonologized down the prosodic hierarchy), and/or pitch reset effects on vocal fold adduction. Because in this paper I focused on glottalization rather than phrase-initial voice quality, further work is needed to determine the exact mechanism responsible for the consistent effect found here. But whatever the reason, it is clear that phrase-initial (abducted) voice quality is a distinct phenomenon from word-initial glottalization (adduction).

4.3 On the cross-linguistic prevalence of word-initial glottalization

In light of the findings on where glottal stops occur and the dissimilarity between glottal stops and phrase-initial voice quality, it is likely that glottal stops are inserted before word-initial vowels that are prominent, especially when they occur in phrase-initial position. Full [ʔ] occurs before prominent initial vowels more often phrase-initially than phrase-medially because of initial strengthening of the glottal adduction gesture. There is no articulatory or acoustic evidence that a glottal stop gesture is always present before word-initial vowels in English (a hypothesis discussed in Pierrehumbert & Talkin (1992), Dilley et al. (1996), and Borroff (2007)). But as mentioned earlier, we cannot disprove this; all we can say is that extreme lenition of a prominence marker is counterintuitive, given the role of prominence and its role in strengthening other segments (Fougeron & Keating 1997, Keating et al. 2003, Cho 2005). Unlike previous researchers who viewed word-initial glottalization as a form of prosodic strengthening more generally (Pierrehumbert & Talkin 1992, Dilley et al. 1996, Borroff 2007), I claim here that the phenomenon is more accurately viewed specifically as a form of prominence strengthening.

Let us now return to the main research question: why then is word-initial glottalization virtually universal across languages (except those with a /ʔ/$\sim$ $\emptyset$ contrast word-initially)? It cannot be because phrase-initial voicing is in general tenser, because our results show that the opposite is true. Moreover, if phrase-initial vocal fold abduction is due to respiratory and/or muscular constraints, such abduction is likely to be universal, and would counter any trend to glottalize phrase-initially. But put the other way round, glottal stops can counteract the effects of phrase-initial vocal fold abduction, which is useful for word-initial vowels. Prominent Utterance-initial vowels must convey prominence despite their low and rapidly-changing subglottal pressure and their increased glottal area. Prominent phrase-initial (but Utterance-medial) vowels must convey prominence despite pitch reset, which likely causes a brief period of glottal abduction. Further, breathy voice quality can be detrimental to tone recoverability (Silverman 1995; 2003), and prominence is usually marked by tones in languages of the world (Jun 2005). The purpose of prominence is to convey salient information (Pierrehumbert & Hirschberg 1990, Ladd 2008), so to indicate prominence on phrase-initial vowels, the speaker must counteract the effects of phrasing on voice quality. A glottal stop ensures that the initial vowel will be produced with non-breathy voice quality. Initial sonorants in prominent syllables do not need to undergo glottalization, because prominence can be conveyed on the following vowel, where the influence of phrase-initial position on the voice quality is weaker. Therefore, under this account word-initial glottalization is not a result of phrase-initial voice quality, but a reaction to phrase-initial voice quality.

Why then would word-initial glottalization rates vary across languages? As noted earlier, there are languages in which word-initial glottalization almost always occurs (e.g., Czech; Bissiri & Volín (2010)); others, like English or German, which show less frequent occurrence; and Spanish or French, where it is thought to be even rarer (Fougeron 2001, Borroff 2007, Bissiri et al. 2011). For languages with very frequent word-initial glottalization, I hypothesize that prominence-induced glottalization has been generalized to all word-initial vowels, even those that are not prominent. As for cases of infrequent word-initial glottalization, I hypothesize that these languages must cue prominence by other means, e.g. through intonation. For example, in Mexican Spanish the most common pitch accent is L+>H* (with a delayed high tone; de-la-Mota et al. (2010)), meaning that the pitch maximum is reached after the stressed syllable. For vowel-initial words (e.g. 'Ana'), this means that the pitch maximum will occur on the second syllable, where the effects of phrase-initial abduction are reduced. Thus, in Spanish cues to prominence extend beyond the stressed syllable. Moreover, post-lexical prominence can be marked by edge tones (e.g., in Korean, Mongolian, and unaccented Japanese), instead of or in addition to marking the head of the prominent word by means of local changes in amplitude, duration, and pitch (Jun 2005). If in a given language prominence is marked by edge tones and not by additional suprasegmental features, one could expect that word-initial glottalization would be rare, because edge-marking prominence on phrase-initial vowels would be adequately conveyed through tones. I leave investigation of the typology of word-initial glottalization to further research.

5 Conclusion

In this paper, I presented two studies on word-initial glottalization. In the first study, I used logistic mixed-effects regression modeling to predict the occurrence of word-initial full glottal stops ([ʔ]) in an English corpus. The results indicated that prominence and phrasing are the most important factors in predicting full glottal stop occurrence. Moreover, prominent word-initial vowels that were not preceded by [ʔ] showed a decrease in H1*-H2*, an acoustic correlate of glottal constriction. Surprisingly, non-prominent phrase-initial vowels did not show signs of glottal constriction. These findings were then confirmed using electroglottographic contact quotient, and extended to Spanish. Based on the results, I proposed and motivated a theory of word-initial glottalization where prominence is the driving force in determining whether word-initial vowels show glottal stop epenthesis, and where higher phrasal domains are responsible for the strength of the glottal stop gesture. Typological differences in the way prominence is cued can account for variable rates of glottalization across languages.

References

Baayen, R. H. (2008). *Analyzing Linguistic Data. A practical introduction to statistics.*. Cambridge: Cambridge University Press.

Bates, D., Maechler, M., & Dai, B. (2008). *lme4: Linear mixed-effects models using S4 classes*. [Computer software program] available from `http://lme4.r-forge.r-project.org/`.

Beckman, M. E., & Ayers Elam, G. (1997). *Guidelines for ToBI labelling, Version 3.*. The Ohio State University Research Foundation.

Bissiri, M. P., Lecumberri, M. L., Cooke, M., & Volín, J. (2011). The role of word-initial glottal stops in recognizing English words. In *Proceedings of Interspeech 2011*. Florence, Italy.

Bissiri, M. P., & Volín, J. (2010). Prosodic structure as a predictor of glottal stops before word-initial vowels in Czech English. In *Proceedings of the 20th Czech-German Workshop on Speech Processing*.

Borroff, M. L. (2007). *A landmark underspecification account of the patterning of glottal stop*. Ph.D. thesis Stony Brook University.

Brunner, J., & Żygis, M. (2011). Why do glottal stops and low vowels like each other? In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 17)* (pp. 376–379).

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990.

Buizza, E., & Plug, L. (2012). Lenition, fortition and the status of plosive affrication: the case of spontaneous RP English /t/. *Phonology*, *29*, 1–38.

Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication*, *15*, 39–54.

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ,i/ in English. *Journal of the Acoustical Society of America*, *117*, 3867–3878.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*, 466–485.

Davidson, L., & Erker, D. (2012). Hiatus resolution in American English: the case against glide insertion. Unpublished manuscript, NYU.

de-la-Mota, C., Butragueño, P. M., & Prieto, P. (2010). Mexican Spanish intonation. In P. Prieto, & P. Roseano (Eds.), *Transcription of Intonation of the Spanish Language* (pp. 319–350). München: Lincom Europa.

DiCanio, C. T. (2009). The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association*, *39*, 162–188.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, *24*, 423–444.

Eddington, D., & Channer, C. (2010). American English has goʔ a loʔ of glottal stops: Social diffusion and linguistic motivation. *American Speech*, *85*, 338–351.

Epstein, M. A. (2002). *Voice quality and prosody in English*. Ph.D. thesis UCLA.

Esling, J. H., Fraser, K. E., & Harris, J. G. (2005). Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics*, *33*, 383–410.

Esposito, C. M. (2012). An acoustic and electroglottographic study of White Hmong phonation. *Journal of Phonetics*, *40*, 466–476.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, *29*, 109–135.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, *101*, 3728–3740.

Fujimura, O. (1990). Methods and goals of speech production research. *Language and Speech*, *33*, 195–258.

Garellek, M. (2012). Glottal stops before word-initial vowels in American English: distribution and acoustic characteristics. *UCLA Working Papers in Phonetics*, *110*, 1–23.

Garellek, M., & Keating, P. (2011). The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association*, *41*, 185–205.

Gordon, M., & Ladefoged, P. (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics*, *29*, 383–406.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: The University of Chicago Press.

Hayes, B. (2009). *Introductory Phonology*. Oxford: Blackwell.

Herbst, C., & Ternström, S. (2006). A comparison of different methods to measure the EGG contact quotient. *Logopedics Phoniatrics Vocology*, *31*, 126–138.

Hirano, M., Ohala, J., & Vennard, W. (1969). The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. *Journal of Speech and Hearing Research*, *12*, 616–628.

Hirose, H. (1997). Investigating the physiology of laryngeal structures. In W. J. Hardcastle, & J. Laver (Eds.), *The Handbook of Phonetic Sciences* (pp. 116–136). Oxford: Blackwell.

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

Howard, D. M. (1995). Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice*, *9*, 163–172.

Hsu, C.-S., & Jun, S.-A. (1998). Prosodic strengthening in Taiwanese: syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics*, *96*, 69–89.

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

Jongenburger, W., & van Heuven, V. J. J. P. (1991). The distribution of (word initial) glottal stop in Dutch. In F. Drijkoningen, & A. van Kemenade (Eds.), *Linguistics in the Netherlands* (pp. 101–110). Amsterdam: John Benjamins.

Jun, S.-A. (2005). Prosodic typology. In S.-A. Jun (Ed.), *Prosodic typology* (pp. 430–458). Oxford: Oxford University Press.

Keating, P. (2006). Phonetic encoding of prosodic structure. In J. Harrington, & M. Tabain (Eds.), *Speech production: Models, phonetic processes, and techniques* Macquarie Monographs in Cognitive Science (pp. 167–186). Psychology Press.

Keating, P., Cho, T., Fougeron, C., & Hsu, C.-S. (2003). Domain-initial articulatory strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Phonetic Interpretation (Papers in Laboratory Phonology 6)* (pp. 143–161). Cambridge: Cambridge University Press.

Kohler, K. J. (1994). Glottal stops and glottalization in German. Data and theory of connected speech processes. *Phonetica*, *51*, 38–51.

Kreiman, J., & Sidtis, D. (2011). *Foundations of Voice Studies*. Oxford: Wiley-Blackwell.

Kuang, J. (2011). *Production and perception of the phonation contrast in Yi*. Master's thesis UCLA.

Ladd, D. R. (1984). Declination: a review and some hypotheses. *Phonology Yearbook*, *1*, 53–74.

Ladd, D. R. (1988). Declination "reset" and the hierarchical organization of utterances. *Journal of the Acoustical Society of America*, *84*, 530–544.

Ladd, D. R. (2008). *Intonational Phonology*. (2nd ed.). Cambridge: Cambridge University Press.

Lombardi, L. (2002). Coronal epenthesis and markedness. *Phonology*, *19*, 219–251.

Mendelsohn, A. H., & Zhang, Z. (2011). Phonation threshold pressure and onset frequency in a two-layer physical model of the vocal folds. *Journal of the Acoustical Society of America*, *130*, 2961–2968.

Milroy, J., Milroy, L., Hartley, S., & Walshaw, D. (1994). Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English. *Language Variation and Change*, *6*, 327–357.

Mompeán, J., & Gómez, F. A. (2011). Hiatus resolution strategies in non-rhotic English: the case of /r/-liaison. In *Proceedings of the International Conference of Phonetic Sciences (ICPhS 17)* (pp. 1414–1417). Hong Kong.

Myers, S. (2012). Final devoicing: Production and perception studies. In T. Borowsky, S. Kawahara, & M. Sugahara (Eds.), *Prosody Matters: Essays in honor of Elisabeth Selkirk*. London: Equinox Publishing.

Orlikoff, R. F. (1991). Assessment of the dynamics of vocal fold contact from the electroglottogram. *Journal of Speech and Hearing Research*, *34*, 1066–1072.

O'Shaughnessy, D., & Allen, J. (1983). Linguistic modality effects on fundamental frequency in speech. *Journal of the Acoustical Society of America*, *74*, 1155–1171.

Ostendorf, M., Price, P. J., & Shattuck-Hufnagel, S. (1995). *The Boston University radio news corpus*. Technical Report ECS-95-001 Boston University.

Pierrehumbert, J. (1995). Prosodic effects on glottal allophones. In O. Fujimura, & M. Hirano (Eds.), *Vocal fold physiology: voice quality control* (pp. 39–60). San Diego: Singular Publishing Group.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. J. Docherty, & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II* (pp. 90–117). Cambridge: Cambridge University Press.

Pierrehumbert, J. B., & Hirschberg, J. (1990). The meaning of intonation contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Plans and Intentions in Communication and Discourse* SDF Benchmark Series in Computational Linguistics (pp. 271–311). Cambridge: MIT Press.

Pompino-Marschall, B., & Żygis, M. (2011). Glottal marking of vowel-initial words in German. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 17)* (pp. 1626–1629). Hong Kong.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, *90*, 2956–2970.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria.

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, *29*, 407–429.

Shue, Y.-L., Chen, G., & Alwan, A. (2010). On the interdependencies between voice quality, glottal gaps, and voice-source related acoustic measures. In *Proceedings of Interspeech 2010* (pp. 34–37).

Shue, Y.-L., Keating, P. A., Vicenik, C., & Yu, K. (2011). VoiceSauce: A program for voice analysis. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS 17)* (pp. 1846–1849). Hong Kong.

Silverman, D. (1995). *Phasing and recoverability*. Ph.D. thesis UCLA. [Published in 1997 by Garland].

Silverman, D. (2003). Pitch discrimination between breathy vs. modal phonation. In J. Local, R. Ogden, & R. Temple (Eds.), *Phonetic Interpretation (Papers in Laboratory Phonology 6)* (pp. 293–304). Cambridge: Cambridge University Press.

Slifka, J. (2000). *Respiratory constraints on speech production at prosodic boundaries*. Ph.D. thesis MIT.

Slifka, J. (2006). Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice*, *20*, 171–186.

Straka, G. (1963). La division des sons du langage en voyelles et consonnes peut-elle être justifiée? *Travaux de linguistique et de littérature, Université de Strasbourg*, *1*, 17–99.

Umeda, N. (1978). Occurrence of glottal stops in fluent speech. *Journal of the Acoustical Society of America*, *64*, 88–94.

Westbury, J. R., & Keating, P. A. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, *22*, 145–166.

Zhang, Z. (2011). Restraining mechanisms in regulating glottal closure during phonation. *Journal of the Acoustical Society of America*, *130*, 4010–4019.