Journal of Phonetics (2001) **29**, 407–429
doi:10.006/jpho.2001.0145
Available online at http://www.idealibrary.com on IDEAL®

Variation in the realization of glottalization in normal speakers*

Laura Redi and Stefanie Shattuck-Hufnagel

Speech Communications Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, MA, U.S.A.

Glottalized voice quality has been observed consistently in normal utterances in a variety of locations, including as an allophone of voiceless stops (in e.g., *Hatfield*, *butler*), in word-initial vowels at intonation phrase onsets and at pitch accents, and at the ends of utterances. In this study of American English, we examine glottalizations at phrase boundaries which are medial or final in an utterance. Tokens are characterized as examples of aperiodicity (irregular periods), creak (lowering of fundamental frequency with near-total damping), diplophonia (alternation in the shape, amplitude, or duration of successive periods) and a fourth category, glottal squeak. Findings include (a) a wide range in the rates of glottalization and in preferred acoustic characteristics across individual speakers, (b) a higher rate of glottalization on words at the ends of utterances than on words at the ends of utterance-medial intonational phrases, and (c) a higher rate of glottalization at the boundaries of full intonational phrases than at intermediate intonational phrases. These patterns will need to be accounted for in any comprehensive treatment of surface phonetic variation in speech.

1. Introduction

The fact that normal speakers exhibit glottalized voice quality during spoken utterances has become a matter of considerable interest to speech scientists in recent years, in part because of growing evidence that this variation serves a communicative function. For example, voice source contrasts can distinguish words and segments in some languages (Ladefoged & Maddieson, 1996; Abramson & Tingsabadh, 2000; Kohler, 2000; Gordon & Ladefoged, this volume), and can serve an allophonic function in others (as in e.g., the glottalization of syllable-final /t/ in many dialects of English). It can also characterize the ends of utterances (Henton & Bladon, 1987), or punctuate the onsets of intonational phrases or pitch accents (Pierrehumbert & Talkin, 1992; Pierrehumbert, 1995; Dilley, Shattuck-Hufnagel & Ostendorf, 1996).

Address correspondence to S. Shattuck-Hufnagel. E-mail: stef@speech.mit.edu
*This work was supported by NIH grants DC02125 and DC00075, NIDCD Training Grant 5-7T32-DC0038, and a NSF Graduate Fellowship.

0095-4470/01/040407 + 23 $35.00/0

One striking aspect of glottalization is its variation across individual speakers, both in its rate and in its acoustic characteristics. This paper describes factors related to this variation, as a contribution to understanding the distribution of glottalization in normal utterances. A better understanding of glottalization will also contribute to an improvement of automatic detection of these regions for speech recognition and analysis, an effort which is already underway (Batliner, Burger, Johne & Kießling, 1993; Hagen, 1997).

Defining glottalization requires careful specification of criteria, and this will be dealt with in detail in the methods section. In acoustic terms, glottalization is understood to be a region in the speech signal characterized by irregularly spaced pitch periods and often accompanied by other characteristics, such as full damping, low f_0, breathiness, or low amplitude (Ladefoged, 1971; Fischer-Jørgensen, 1989; Klatt & Klatt, 1990; Pierrehumbert & Talkin, 1992). These characteristics are recognized as lending a perceptual impression of a glottal gesture or disturbance in modal voice quality (Rozsypal & Millar, 1979; Hillenbrand & Houde, 1996; Pierrehumbert & Frisch, 1997). In the following discussion, we summarize factors that have been proposed to affect rates and acoustic manifestations of glottalization. While some of these factors have been studied carefully, others have received less attention. The current study therefore aims to supplement the existing literature by examining the effects of some factors in greater depth than has been reported elsewhere. This will be accomplished by controlling for text, professionalism of the speaker, and position in the utterance and intonational phrase.

1.1. *Variation with phrasal position*

The rate of glottalization varies according to position within utterances and within intonational structure. For example, Henton & Bladon (1987) found that creak was more likely in syllables at the ends of utterances than elsewhere. In addition, intonational phrase structure has been shown to influence glottalization rates for word-initial vowels. When Pierrehumbert & Talkin (1992) examined the rates of glottalization of word-initial vowels for two speakers, they found more glottalization at the onset of intonational phrases than at other locations, at least for reduced vowels. Pierrehumbert (1995) also reported more glottalization for word-initial vowels at pitch accents. Dilley *et al.* (1996) extended these observations to five professional radio news readers. They found similar results, showing that glottalization of word-initial vowels was more likely at intonation-phrase onsets and at pitch accents than at other locations. Moreover, they found that the pattern for reduced vowels distinguished between two levels of intonational phrase, full *vs.* intermediate.

In sum, it appears that glottalization is statistically more likely (although not inevitable) at prosodically significant locations such as phrase boundaries, utterance boundaries and pitch accents, at least in English. The likelihood of glottalization on lexical items at the ends of utterance-medial phrases has not been reported prior to the current study.

1.2. *Variation with gender*

Data from the literature conflict with regard to whether males or females glottalize more frequently. For two dialects of British English (Received Pronunciation and "Modified Northern"), Henton & Bladon (1987) reported that their male speakers glottalized utterance-final syllables much more often than their female speakers did. Rates of glottalization in utterance-final syllables for speakers using the Received Pronunciation dialect were 65% for males *vs.* 35% for females, while the rates for speakers who used the "Modified Northern" dialect were 80% for males *vs.* 27% for females. In contrast with these findings, several other studies have found higher rates of glottalization for female speakers than males. Dilley *et al.* (1996) observed that female radio news readers in their study of American English glottalized word-initial vowels at higher rates (40%, 44% and 38%) than the male speakers (24% and 13%). Similarly, Huber (1988) found that the females in his study of Swedish glottalized at rates almost twice those of the males, in several contexts.[1] Byrd (1994) also found that females produced more glottalization in a variety of phonological contexts than males did.[2]

In sum, studies looking at different dialects, different speakers, and different locations in the utterance have produced conflicting results about whether males or females glottalize more frequently. The factors which contribute to gender differences in rate of glottalization may be anatomical, sociolinguistic, structural, or perhaps a combination. In this study, we examine the effects of gender while controlling for texts, position in utterance and level of professionalism.

1.3. *Variation with segmental context*

Pierrehumbert (1995) looked at the likelihood of glottalization of syllable- and word-final voiceless stops in American English, and found that the nature of the following consonant has a significant effect on the glottalization rate. Dilley *et al.* (1996) studied the effect of preceding context on the glottalization of word-initial vowels in their radio news corpus. They found that glottalization was significantly more likely after a measurable pause (although it also occurred significantly often after intonational phrase boundaries without a pause, and pauses tended to occur at intonational phrase boundaries in this corpus because of the low rate of processing-related disfluencies). They also found that in non-phrase-initial position, the only preceding segment context that made a significant difference in glottalization rate was the class of vowel segments. This is consistent with other studies which report a high rate of word-initial-vowel glottalization at a vowel–vowel hiatus (Umeda, 1978; Gimson, 1989; Pierrehumbert & Frisch, 1997). It is our belief, based on results in Dilley *et al.* (1996) and Pierrehumbert (1995), that glottalization at vowel–vowel hiatus is correlated with the occurrence of a pitch accent on the second vowel, but the study to be reported here does not investigate the influence of preceding segmental context.

1.4. *Variation with dialect*

Within a given language, different dialects have been observed to make use of glottalization in different contexts to varying degrees. This dialectic variation may be either allophonic or pragmatic in nature. For example, there is variation across different

[1] Huber's criterion for glottalization in word-initial vowels appears to have been a conservative one: a single pulse of irregularity. It is unclear what effect this stringent criterion may have had on his results.

[2] The speech materials used in the study by Byrd (1994) came from the TIMIT database (Zue & Seneff, 1988), which includes a time-aligned phonetic transcription. A marker for glottal stop was included in phonetic transcriptions of this database, and Byrd found that the relative frequency of this glottal stop marker was higher for females than males. In reporting that Byrd's results indicated a higher rate of glottalization for females than males, we are inferring that the glottal stop marker used in the TIMIT corpus was indicated in cases for which we would have identified as glottalization, as defined elsewhere in this paper.

dialects of English with regard to the frequency with which glottalized or nonglottalized variants of the voiceless stop /t/ are realized, the environment in which they occur, and the extent to which stops other than /t/ are affected (Kerswill, 1987; Docherty & Foulkes, 1995). To illustrate, speakers of General American English are more likely to produce flapped /t/ in positions where speakers of many varieties of British English glottalize. Another example of dialect differences in allophonic variation is the use of glottalization to reinforce or replace final /d/ in stressed syllables in African American English (AAE) (Fasold, 1981; Kohl & Anderson, 2000). (Devoicing of these glottalized stops also takes place, so that speakers of AAE might produce, for example, the lexical monosyllable *bed* as [bɛːʔt].) In contrast, final stop glottalization is commonly described to be limited to voiceless /p,t,k/ in other varieties of American English. Moreover, Byrd (1994) reports that speakers of American English from the north and south regions exhibit more frequent use of glottal stop than speakers from other regions of the country.

Some authors have also described differences in the use of glottalization for pragmatic reasons. For example, Laver (1980) and others have suggested that creak may be used to signal the end of a speaker's turn. Other authors have described differences among dialects in this use of glottalization. For example, glottalization may mark the end of a turn in London Jamaican (Local, Wells & Sebba, 1985), but it is avoided in turn-final position in the Tyneside dialect (Local, Kelly & Wells, 1986). Finally, as noted earlier, dialect appears to interact with gender with respect to glottalization rates (cf. Henton & Bladon, 1987; Byrd, 1994; Dilley *et al*., 1996).

1.5. *Variation with individual speaker*

1.5.1. *Variation in glottalization rate across speakers*

A few studies have shown that different individuals glottalize at very different rates. For example, rate of glottalization for word-initial vowels produced by five English-speaking professional radio announcers who each produced different texts varied from 13% to 44% (Dilley *et al*., 1996). For four nonprofessional speakers and five professional speakers who all produced the same text, the rates ranged from less than 1% to 34% (Dilley & Shattuck-Hufnagel, 1995). Huber (1988) found that four Swedish speakers glottalized the conjunction *och* at rates ranging from 12% to 71% in coordinate-clause-initial positions in read speech. Moreover, although Henton & Bladon (1987) did not systematically study inter-speaker variation in their corpus of British English, they note that a few of their speakers spoke "with almost continuous creak"; nine of the 10 speakers who produced more than 65% creaky syllables were male speakers of the "Modified Northern" dialect.

In sum, it appears that individual speakers differ substantially in their glottalization rates at a variety of locations within the utterance. As is the case for gender effects, the factors which are responsible for these differences are not yet clear. In this study, we examine differences in rate of glottalization for individual speakers, while controlling for other factors.

1.5.2. *Variation in preferred acoustic characteristics across speakers*

Many investigators have reported that the acoustic characteristics of waveforms that are perceived as glottalized can vary substantially from utterance to utterance, or even within utterances. These variations have been particularly well-studied for allophonic variation. For example, Kohler (1994) examined phonetic variability in morphologically-conditioned productions of /ʔ/ in German, in a corpus of 25 male and 25 female talkers. He found that what is traditionally referred to as "glottal stop" was realized with variable phonetic characteristics: 27% of locations involved several irregular pitch periods without the sustained period of silence normally associated with a stop; 27% involved a sustained silence plus irregular pitch periods; and only 15% of locations showed the single irregular pulse with silence which is often discussed as typical of glottal stop. Fully 21% of the locations where glottal stop was expected showed no evidence of nonmodal phonation. Similarly, Docherty & Foulkes (1995, 1999) found for the Tyneside dialect of British English that voiceless stops heard as glottalized were rarely produced with a period of silence characteristic of a glottal stop, and were more frequently produced with continuous but irregular voicing. Finally, Fischer-Jørgensen (1989) found a wide range of variation in acoustic realization of the *stød* in standard Danish, both across speakers and within the same speaker on different occasions.

A wide range of variation in acoustic shape has also been reported for phrasally-governed glottalization episodes in natural speech. For example, Dilley & Shattuck-Hufnagel (1995) reported correlates such as diplophonia (alternation in shape, amplitude, or duration of successive pulses), low amplitude, low fundamental frequency, near-total damping and breathiness in some regions of glottalization in word-initial vowels. In addition, they noted that different speakers showed different preferences among these acoustic characteristics. In fact, one speaker showed a strong tendency to produce reduced amplitude in prosodically significant locations where other speakers glottalized.

In order to investigate the factors that influence these acoustic differences, many researchers have found it helpful to develop categories of glottalization events. While perceptual criteria have been shown to be useful in informing acoustic analyses and categorization (see Gerratt & Kreiman, this volume), the method of categorization described here relies jointly on perceptual and acoustic criteria. However, at least some categorization schemes have been proposed based solely on observable acoustic characteristics of glottalized regions. For example, Huber (1988) proposed the categories of (i) creaky voice, or period-to-period irregularity, (ii) creak, or sustained low frequency accompanied by almost total damping of individual pitch periods, (iii) diplophonia or alternation in the shape, amplitude or period duration of successive pulses, and (iv) glottal stop, or single pulses of irregularity. Using this system to analyze utterances from two male and two female speakers of Swedish, Huber found that speakers of both genders exhibited each of the types of glottalization, but different speakers showed different distributions. For example, the two female speakers produced more creaky voice than creak, while the two males showed the opposite tendency. Other categorization systems have proposed a number of dimensions along which a glottalized event can vary quasi-independently. For example, Batliner *et al.* (1993) developed a classification system known as Münchner Schema für Laryngalisierungs-Identification (MÜSLI) which they used to code each glottalized token along six acoustic dimensions. These dimensions included number of pitch periods in the glottalized region, degree of damping, degree of amplitude difference with respect to context, f_0 difference with respect to context, and f_0 difference within the glottalized region. Two phoneticians made judgments of these characteristics for glottalized regions in a portion of the SPONTAN database of spoken German, using both auditory criteria and examination of visual displays of the speech signal. They agreed on the value for all six dimensions in 81% of cases.

In summary, earlier studies of glottalization have noted a tendency toward wide variation in rate of occurrence and in preferred acoustic characteristics across languages, dialects, individual speakers and phrasal position. One sparsely-studied area is individual speaker variation in the rate and acoustic shape of phrase-final glottalization, particularly for utterance-medial phrases. The aim of this study is to compare the rates and characteristics of glottalization in regions where phrase boundaries are likely, for a sample of speech from both professional and nonprofessional speakers of American English. The results will provide a step toward the goal of understanding the governing factors and articulatory mechanisms involved in this phenomenon.

2. Method

2.1. *Speech materials*

In selecting speech materials for study, we chose materials for which as many factors as possible were controlled, so as to maximize the likelihood of ascertaining the effects only of relevant factors under study. The speech materials ultimately selected were controlled for text, and therefore for segmental context (for both the Labnews and ABC corpora, described below), for position within the phrase and/or utterance (again, for both corpora) and for prosodic location (for the Labnews corpus). Given these constraints, perceptual and acoustic evidence of glottalization was ascertained for a subset of lexical items which met the criteria described below. Glottalization could occur at any position within the lexical item.

Speech materials used in this study came from two corpora of spoken utterances. The first corpus, which we call Labnews, consisted of read speech which had been produced by six professional, American radio news announcers (female speakers F1, F2, and F3 and male speakers M1, M2, and M3). A subset of these materials had also been produced by four nonprofessional speakers (female speakers FJ and FS, and male speakers MK and MM). Among the nonprofessional speakers, FJ, FS, and MM were American; MK was a Canadian who had been living in the United States for more than 20 years.

Since the majority of speech had earlier been hand-labeled for intonational phrase boundaries and pitch accents using the ToBI system (Beckman & Pierrehumbert, 1986; Beckman & Ayers-Elam, 1997), it was possible to directly compare intonational phrase boundary information with glottalization rate and distribution. In particular, all of the speech produced by the professional speakers had been previously annotated for prosodic structure and prominence, and half of the speech produced by nonprofessionals had been labeled.

The second corpus, which we call the ABC corpus, consisted of a set of read sentences produced by each of four nonprofessional speakers of American English (female speakers HH, RM, and SK, and male speaker ES). In contrast to the materials in the Labnews corpus, material in the ABC corpus had not been labeled for prosody. However, informal listening confirmed that all four speakers habitually placed a phrase boundary at the location of the comma and at the end of the sentence.

Lexical items selected for this study occupied a subset of locations where glottalization was thought likely to occur. One probable location for glottalization to occur is at the end of a large prosodic phrase boundary (also known as a full intonational phrase

TABLE I. Number of lexical items at utterance-medial full and intermediate intonational phrases examined for evidence of glottalization for six speakers

Speaker	No. at full intonational phrases	No. at intermediate intonational phrases	Total
F2	97	33	130
M1	65	55	120
FJ	48	60	108
FS	71	46	117
MK	65	42	107
MM	64	43	107

boundary). Therefore, for the professional and nonprofessional speakers in the Labnews corpus, those lexical items for which all speakers in the speaker group (professional or nonprofessional) had produced a full intonational phrase boundary were examined for glottalization.[3] Restricting analysis to locations where a large prosodic boundary was present provided a control for the effects of prosodic boundaries. A total of 78 lexical items per speaker for each of the six professional speakers were identified in this manner, and 56 lexical items per speaker for each of the nonprofessional speakers were examined for glottalization, yielding 468 tokens total for professional speakers and 224 tokens total for nonprofessional speakers.

The Labnews corpus also permitted a comparison between glottalization rates at full intonational phrase boundaries and at intermediate intonational phrase boundaries, for a restricted set of materials and speakers. For a single radio news story produced by six of the speakers (professional speakers F2 and M1; nonprofessional speakers FJ, FS, MK, and MM), the locations of words at the ends of utterance-medial full and intermediate intonational phrases were determined. These items were subsequently examined for evidence of glottalization. Table I indicates the number of lexical items examined at utterance-medial full and intermediate intonational phrases for each speaker.

For the ABC corpus, examination of the speech signal was again restricted to locations where glottalization was thought likely to occur, i.e., at locations where intonational phrase boundaries were thought likely. The corpus consisted of 10 sentences of the form *Please say* (*word* A) *or* (*word* B) *and* (*word* C) *will stay/play*. Each sentence was presented in two stimulus formats: either with a comma after word A (e.g., *Please say Charlestown, or Brookline and Woodstock will play*). Words A, B, and C were proper names (e.g., *Charlestown*, *Brookline*, *Woodstock*; *Tom*, *Bob*, *Jim*); the final verb *play* was used with nouns appropriate for athletic teams and the verb *stay* for names of individuals (Shattuck-Hufnagel & Turk, 1998; Turk, 1999). Each sentence was produced in each constituent format by each speaker, yielding 20 sentences (10 sentence pairs). Four lexical items were selected for study from each sentence because they occurred at likely pre- or

[3]The four nonprofessional Labnews speakers had produced only two of the four radio news stories, and of these, one had been prosodically labeled on a prior occasion. For the second story which had not been labeled for prosody earlier, lexical items were selected for study for which all six of the professional news speakers had produced a full intonational phrase boundary. *Post hoc* auditory analysis confirmed that each of the four nonprofessional speakers had produced a full intonational phrase boundary at these locations as well.

postboundary locations. For the boundary-after-A versions, the four lexical items selected were: *say* (a potential boundary location), the preboundary proper name (A), the postboundary conjunction *or*, and the final verb (*play* or *stay*). For the boundary-after-B versions, the four lexical items selected for study were: *say*, the preboundary proper name (B), the postboundary conjunction *and*, and the final verb (*play* or *stay*). A total of 80 tokens per speaker from the ABC corpus were examined, yielding 320 tokens total.

2.2. *Classification of glottalized events*

For each lexical item examined for evidence of glottalization, the following determinations were made: (1) whether or not the lexical item presented *both* acoustic and auditory evidence of glottalization; and (2) if acoustic evidence of glottalization was present, what sort of irregularit(ies) were present, according to the set of categories described below. A detailed description of what counted as either acoustic or perceptual evidence of glottalization is offered later in this section. In all cases, visual inspection of the acoustic speech waveform was the primary means of assessing acoustic evidence of glottalization, but spectrograms and/or fundamental frequency contours were sometimes consulted as well. Visual displays were made using XWaves software from Entropic Corporation.

To aid in investigation of variability in acoustic characteristics of glottalization, four acoustic categories of glottalization were developed, and a given region of glottalization could display one or all of these characteristics. The four categories were: (1) irregularity in duration of glottal pulses from period to period which we term *aperiodicity* (Fig. 1); (2) prolonged low fundamental frequency accompanied by almost total damping of glottal pulses or *creak* (Fig. 2); (3) regular alternation in shape, duration, or amplitude of glottal periods or *diplophonia* (Fig. 3); and, finally, (4) *glottal squeak*, or a sudden shift to relatively high sustained f_0, which was usually very low amplitude (Fig. 4).[4,5]

The first three acoustic correlates appear to correspond well to the categories of creaky voice, creak, and diplophonia, respectively, determined by Huber (1988), and were arrived at independently. The fourth characteristic, glottal squeak, represents what we believe to be a previously undescribed type of nonmodal phonation in normal speakers; it was produced relatively rarely in our corpus, and only by certain speakers in our sample.

In order to be counted as glottalized, a candidate word was required to exhibit both perceptual and acoustic evidence of glottalization. Perceptual and acoustic evidence were sought concomitantly for each token. Perceptual correlates of glottalization included salient auditory impressions of a glottal gesture, roughness, or creakiness. There were some perceptual differences among the acoustic categories of glottalization,

[4]The purpose of the figures is to illustrate categories of glottalization used in the present study. A few examples were taken from a larger corpus of glottalized events which includes as a subset the material used in the current study.

[5]Irregular vibration may arise even when the glottis is configured as for modal vibration, if other conditions for stable modal phonation are not met, e.g., if the trans-glottal pressure difference is not appropriate. Thus, the criteria used for glottalization in this study, i.e., auditory and visual evidence for irregular pitch periods, define a category of events which may include some tokens where the irregular vibration arises from configurations other than sudden abduction or adduction of the vocal folds. See Section 2.3 in Hanson, Stevens, Kuo, Chen & Slifka (this volume) for discussion.

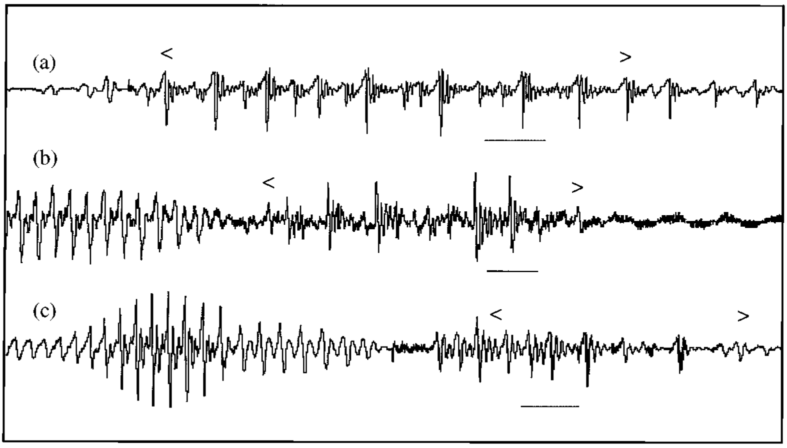

Figure 1. Glottalized tokens illustrating aperiodicity. Angled brackets above each waveform indicate the region of aperiodicity, and horizontal bar indicates 20 ms. For (a) the depicted utterance is *law*. In (b) the utterance is (*me*)*mber*. For (c) the utterance is (*se*)*venty*. Speakers are F2, FJ, and F1, respectively.

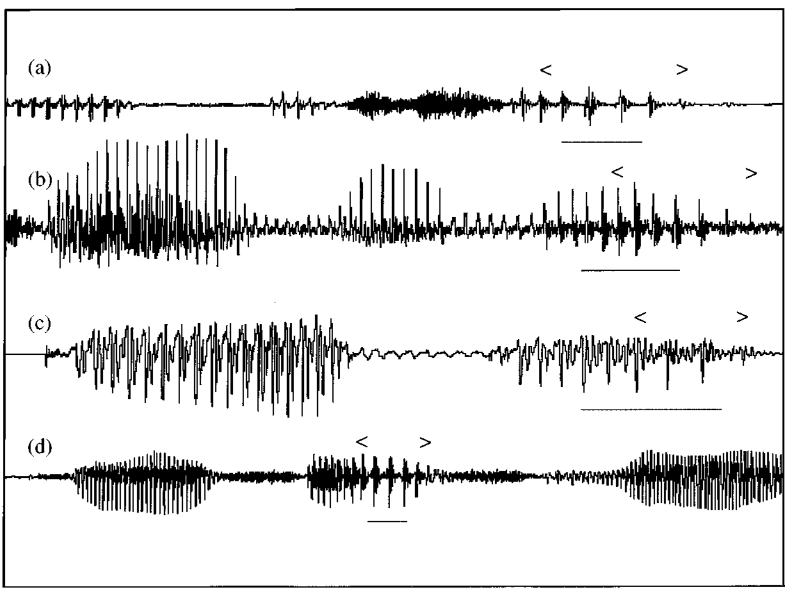

Figure 2. Glottalized tokens illustrating creak. Angled brackets above each waveform indicate the region of creak, and horizontal bar represents 50 ms. For (a) the depicted utterance is *officer*. In (b) the utterance is *governor*. For (c) the utterance is *doubled*. For (d) the utterance is *process#wor*(*ship*). Speakers are F1, FS, M3, and F2, respectively.

consistent with findings that listeners can reliably distinguish some categories of acoustic irregularity (Gerratt & Kreiman, this volume). For aperiodicity and creak, perceptual characteristics included an impression of either consonant-like abruptness and/or general lowness of pitch, but note that for these cases, no distinct, identifiable pitch could be

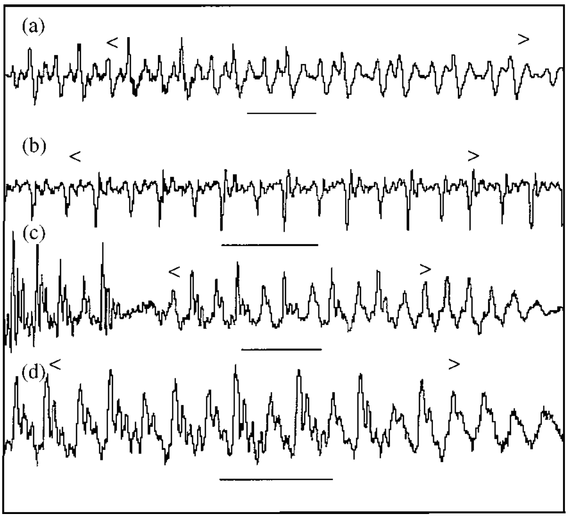

Figure 3. Glottalized tokens illustrating diplophonia. Angled brackets above waveforms indicate the diplophonic region, and the horizontal bar indicates 20 ms. For (a) the depicted utterance is (*A*)*mer*(*ican*). For (b) the utterance is (*surveillan*)*ce*#*i*(*n*). For (c) and (d) the utterance is (*p*)*lay*. The speaker for (a) and (b) is F2, for (c) the speaker is HH, and for (d) the speaker is RM.

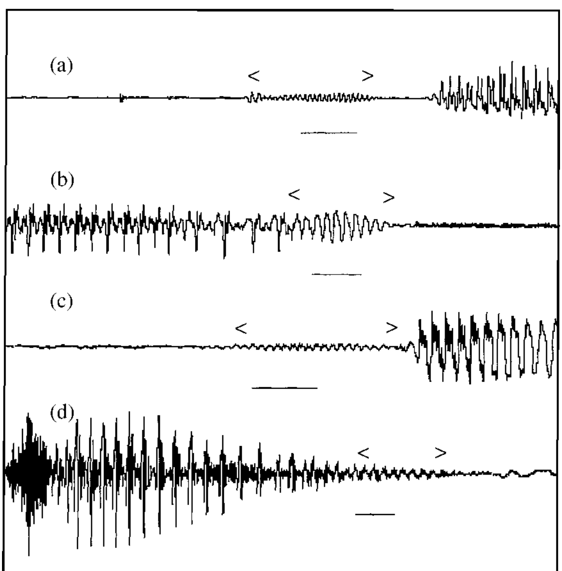

Figure 4. Glottalized tokens illustrating glottal squeak. Angled brackets above each waveform indicate the glottal squeak, and the horizontal bar indicates 25 ms. For (a) the utterance is *or*#*R*(*angoon*). In (b) the utterance is *launched*. For (c) the utterance is (*McWalden*)#*an*(*d*). For (d) the utterance is (*sinis*)*ter*. Speakers are HH, F2, SK, and F3, respectively.

discerned. For diplophonia, the perceptual correlate was a rough voice quality which was usually accompanied by the percept of a distinct pitch approximately an octave below the pitch perceived for a nearby modal region, which is consistent with period doubling. Glottal squeak almost always occurred adjacent to other types of glottalization, and was perceived as an instantaneous shift to a relatively high-pitched, low amplitude voice quality. One author performed the initial classification of tokens, and those tokens for which acoustic and/or perceptual evidence of glottalization was weak were examined by both authors in consultation. Tokens were discarded for which both listeners did not agree that both acoustic and perceptual evidence of glottalization was present.

If a token exhibited perceptual evidence of glottalization, acoustic evidence was also sought. To qualify as exhibiting aperiodicity, we required a token to present visual evidence of successive pitch periods for which incremental changes in duration were discontinuous. For example, a region might display a jump from a relatively shorter pitch period to a longer one, then back to a short period, or it might exhibit the converse—a sequence of long-short-long periods. In contrast, a region which evidenced a gradual decrease in pitch period ending in a very low f_0 would not have been aperiodic in our terms, but rather an example of creak (see next paragraph). Differences in the duration of the longest pitch period of the aperiodic region compared with pitch period durations in the nearest modal region were almost always greater than 2 ms. This should be well above the threshold for detectability of period-to-period durational variations (Pollack, 1971).

Tokens which were counted as exhibiting creak typically displayed gradual widening in pitch period, resulting in a very low fundamental frequency with associated strong damping of pitch periods. Alternatively, a few tokens in intervocalic position exhibited a dip in fundamental frequency. The unifying characteristic of these two sorts of acoustic phenomena was a decrease in fundamental frequency (followed by an increase, in the case of intervocalic creak). A fundamental frequency drop has been shown to be a sufficient cue to elicit perception of glottal stop (Hillenbrand & Houde, 1996; Pierrehumbert & Frisch, 1997) and has been observed previously in locations where glottalization is also likely, such as the onset of vowel-initial works (Dilley *et al*., 1996).

Next, the assessment of diplophonia involved visual identification of regions of alternation of pitch periods (in a simple repeating pattern or a more complex pattern) which had different amplitudes, shapes, or period lengths.

Finally, glottal squeak was characterized by an instantaneous increase in fundamental frequency which was subsequently sustained for multiple periods. Squeak was usually produced with low amplitude. We theorize that it occurs as a result of a shift from either modal or vocal fry register to falsetto register (Hollien, 1974); glottal squeaks were rare in our corpus and were almost always accompanied by other acoustic evidence of glottalization in adjacent regions.

We note that while the method employed in this study for assessing acoustic evidence of glottalization was visual inspection of the speech waveform, other measures have sometimes been used. For example, measurements of the relative amplitudes of $H1$ and $H2$ have been used as a diagnostic for glottalization, as well as relative amplitude of $H1$ and $F1$ (e.g., Kirk, Ladefoged & Ladefoged, 1984; Klatt & Klatt, 1990). While time did not permit us to use this broader range of measures, our results suggest that analysis of the acoustic waveform does provide a highly reliable diagnostic of glottalization.

TABLE II. Overall rates of glottalization for the 14 speakers in this study

Speaker	Rate	Speaker	Rate
F1	68	FS	55
F2	88	MK	52
F3	64	MM	80
M1	13	ES	73
M2	39	HH	49
M3	47	RM	30
FJ	51	SK	41

3. Results

3.1. *Individual variation in rate*

The overall rates of glottalization for all speakers are given in Table II. The rates are broken down further in Figs 5 and 6 into overall rate, rate in phrase-final but utterance-medial position, and rate in utterance-final position. Fewer than 2% of tokens were discarded based on uncertainty about their status as glottalized or not. The statistical tests described are tests of proportions based on the standard normal distribution. In each case, the value of p given is the least significant value for the set of comparisons described, though in some cases, this value was more significant for individual comparisons.

The rate of glottalization differed dramatically among the Labnews speakers. For the professional speakers, the overall rates ranged from 88% (female speaker F2) to 13% (male speaker M1). Glottalization rates for all six speakers were significantly different from one another in pairwise comparisons except F1 *vs.* F3 ($p < 0.01$, $z \geq 2.605$ for all pairwise comparisons except F3 *vs.* M3, which was significant at $p < 0.05$; $z = 2.107$). For the nonprofessional speakers, the rates ranged from 80% (male speaker MM) to 51% (female speaker FJ). Speaker MM glottalized at a rate which was significantly different from all three other speakers ($p < 0.005$; $z \geq 2.833$); no other pairwise comparisons were significant.

For the nonprofessional speakers in the ABC corpus, the rates ranged from 73% (male speaker ES) to 30% (female speaker RM) (Fig. 6). All speakers from this corpus glottalized at significantly different rates in pairwise comparisons ($p < 0.01$, $z \geq 2.480$), except HH *vs.* SK (49 *vs.* 41%) and RM *vs.* SK (30 *vs.* 41%), which were not significantly different.

3.2. *Position within the utterance*

Figs 5 and 6 show that for all 14 speakers, rates of glottalization in utterance-final position were invariably higher than rates in phrase-final but utterance-medial position. These comparisons were significant for four of the 10 Labnews speakers ($p < 0.01$, $z \geq 2.44$ for F1 and M2; $p < 0.02$, $z = 2.197$ for F3; and $p < 0.05$, $z = 1.922$ for MK).

For the ABC corpus, it was discovered *post hoc* that speakers had placed intonational phrase boundaries consistently at all proper noun lexical items (words A and B) utterance-medially, but they had not consistently placed intonational phrase boundaries at the word *say*. Fig. 6 therefore shows a comparison between rates of glottalization on

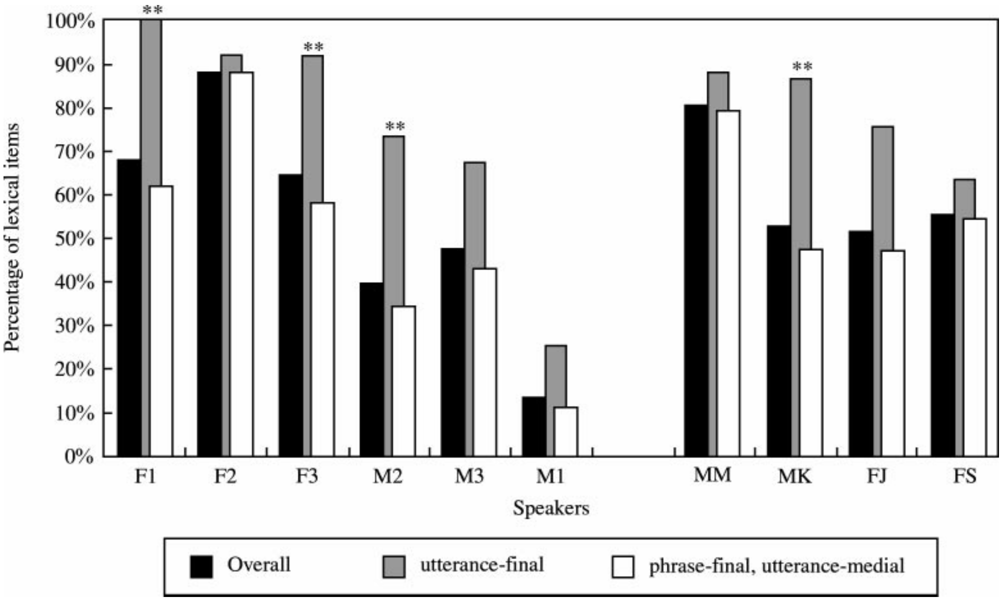

Figure 5. Rates of glottalization for professional speakers (F1, F2, F3, M1, M2, M3) and nonprofessional speakers (FJ, FS, MK, MM), all of whom read the same text from the Labnews corpus. Rates given are overall rate, utterance-final rate, and phrase-final, utterance-medial rate. Rates of utterance-final glottalization were significantly higher than rates of phrase-final but utterance-medial glottalization at $p < 0.025$ or less for some speakers as indicated by **.

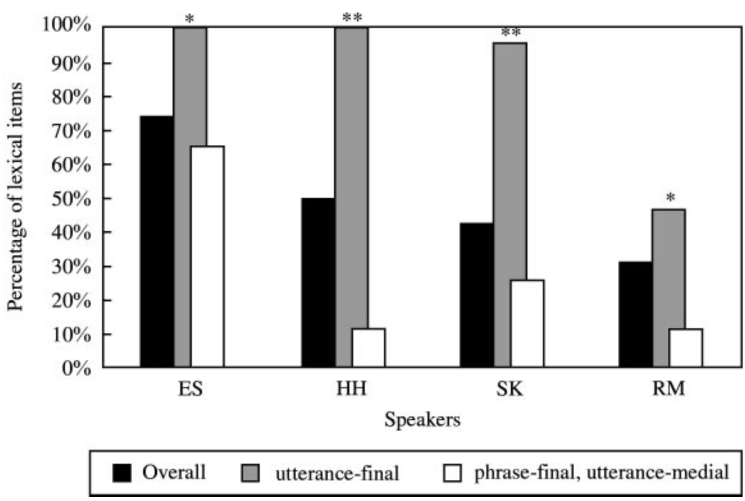

Figure 6. Rates of glottalization for several nonprofessional speakers (females HH, RM, and SK, and male ES) in the ABC corpus. Rates shown are overall rate, utterance-final rate, and phrase-final but utterance-medial rate. Rates of utterance-final glottalization were significantly higher than rates of phrase-final glottalization for all speakers. ** = $p < 0.001$, * = $p < 0.01$.

utterance-final tokens and rates on phrase-final but utterance-medial proper noun tokens only (i.e., not including tokens of *say*). The figure shows that all speakers produced significantly more glottalization in the utterance-final position than in phrase-final but utterance-medial position ($p < 0.001$, $z \geq 4.518$ for HH and SK; $p < 0.01$, $z \geq 2.479$ for ES and RM).

3.3. *Individual variation in preferred acoustic characteristics*

We found significant differences in the rate with which speakers produced different types of glottal irregularity. Figs 7 and 8 show the percentage of glottalized tokens for each speaker which manifested the different acoustic characteristics examined in this study. It is clear that speakers utilized the various characteristics at different rates, with some speakers apparently preferring certain characteristics over others.

3.3.1. *Aperiodicity vs. creaky voice*

Figs 9 and 10 show the number of glottalized tokens which exhibited aperiodicity and creak for the professional and the nonprofessional speakers, respectively. Five of the six professional speakers produced aperiodicity more frequently than creak, and three of the six comparisons were significant (two-sided test; $p < 0.001$, $z \geq 3.875$ for F2 and M2; $p < 0.05$, $z = 2.021$ for F3). Among the nonprofessional speakers, there was also a tendency to produce aperiodicity more frequently than creak. Seven of the eight speakers produced aperiodicity more frequently than creak, and four of the eight comparisons were significant ($p < 0.001$, $z \geq 4.683$ for HH and ES; $p < 0.01$, $z \geq 2.679$ for FJ and MM).

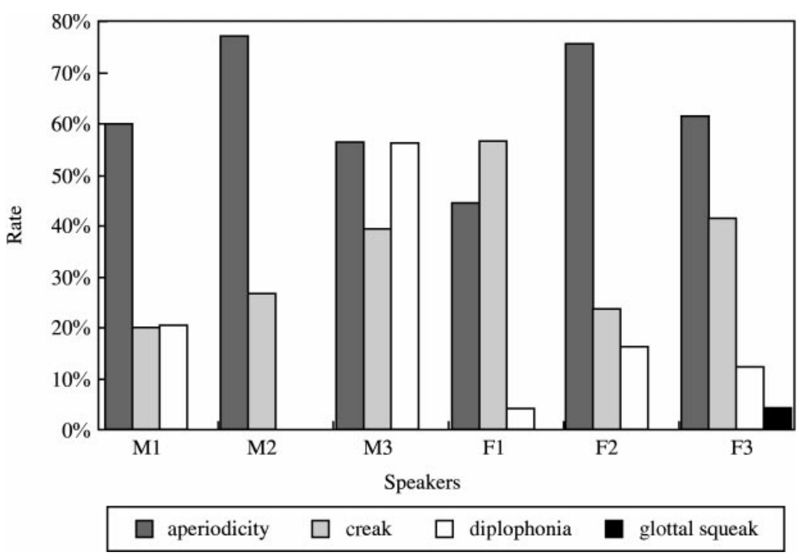

Figure 7. Percentage of glottalized tokens exhibiting aperiodicity, creak, diplophonia, or glottal squeak for six professional speakers.

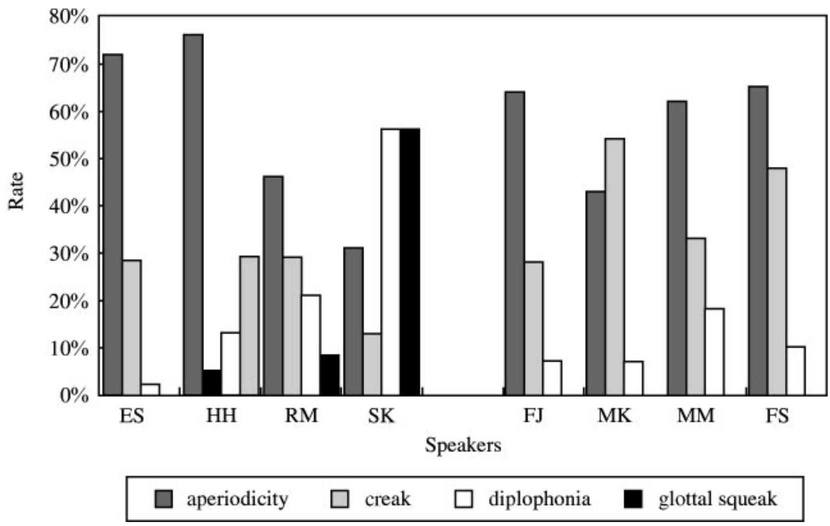

Figure 8. Percentage of glottalized tokens exhibiting aperiodicity, creak, diplophonia, or glottal squeak for eight nonprofessional speakers. Speakers ES, HH, RM and SK were from the ABC corpus, while speakers FJ, MK, MM, and FS were from the Labnews corpus.

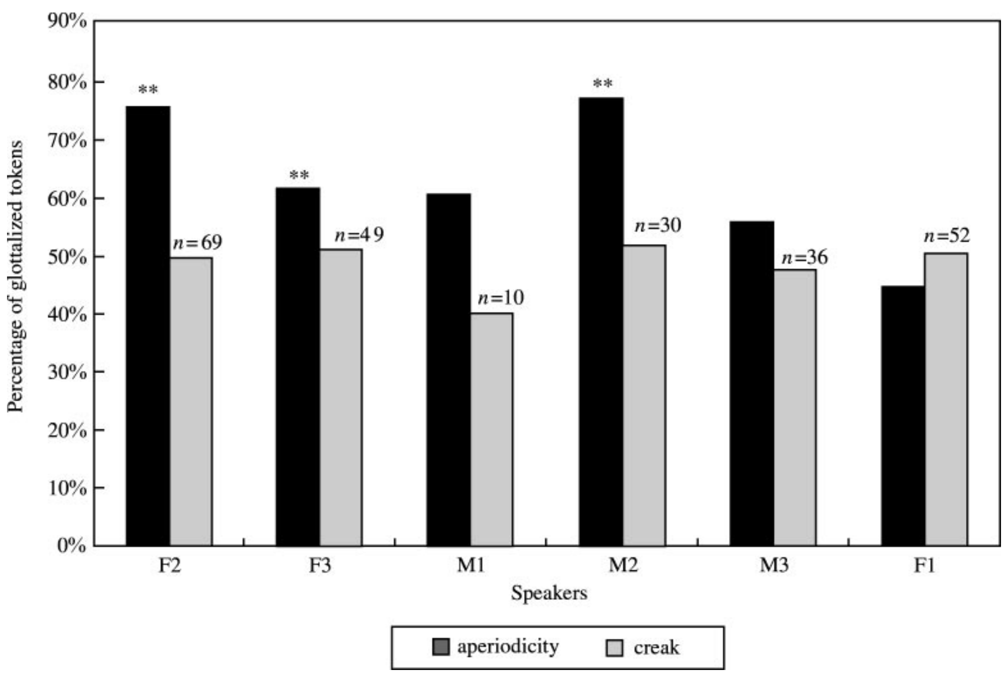

Figure 9. Percentage of glottalized tokens exhibiting aperiodicity or creak for six professional speakers. ** = $p < 0.001$; * = $p < 0.05$.

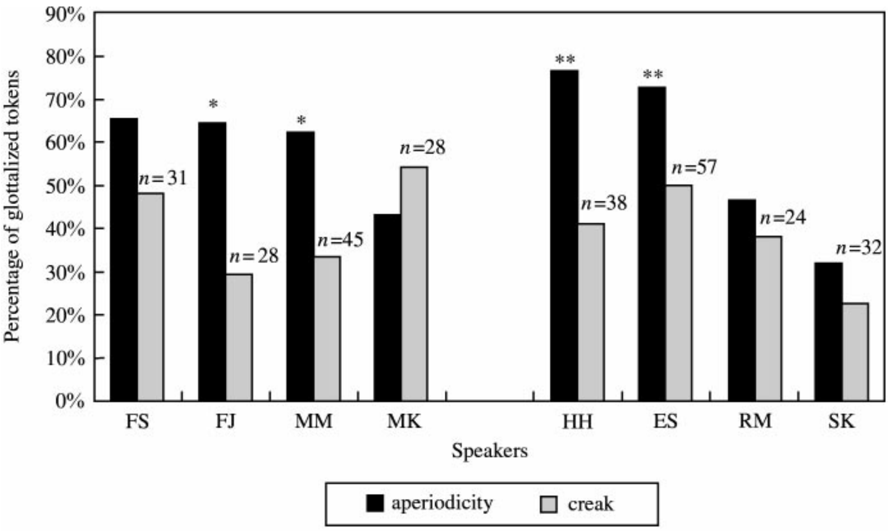

Figure 10. Percentage of glottalized tokens exhibiting aperiodicity or creak for eight nonprofessional speakers. Speakers FS, FJ, MM, and MK were from the Labnews corpus, while speakers HH, ES, RM, and SK were from the ABC corpus, while speakers. ** = $p \leq 0.001$; * = $p < 0.01$.

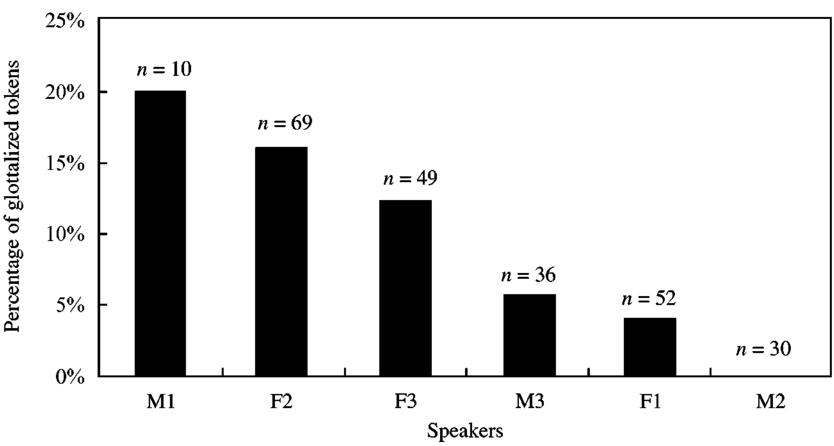

Figure 11. Percentage of glottalized tokens exhibiting diplophonia for professional speakers.

3.3.2. *Diplophonia*

Fig. 11 shows the proportion of glottalized tokens that exhibited diplophonia for the professional speakers, and Fig. 12 gives the same result for the nonprofessional speakers. Among the professional speakers, speaker F2 produced glottalization in the form of diplophonia significantly more often than either F1 or M2 (two-sided test; $p < 0.05$,

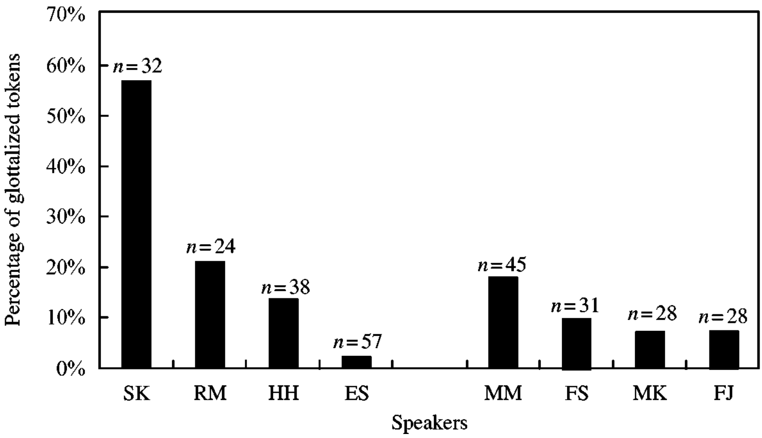

Figure 12. Percentage of glottalized tokens exhibiting diplophonia for nonprofessional speakers. Speakers SK, RM, HH, and ES were from the ABC corpus, while speakers MM, FS, MK, and FJ were from the Labnews corpus.

$z \geq 2.127$). Moreover, speaker F3 produced more diplophonia than speaker M2 ($p < 0.05$, $z \geq 1.994$). For the nonprofessional Labnews speakers, no comparisons were significant. For the nonprofessional speakers in the ABC corpus, ES produced diplophonia the least frequently, and all speakers produced diplophonia significantly more often than ES ($p < 0.05$, $z \geq 2.239$). SK produced diplophonia the most frequently, producing more instances than any of the other nonprofessional speakers ($p < 0.01$; $z \geq 2.666$).

3.3.3. *Glottal squeak*

Glottal squeak was quite rare in our corpus. Of 529 glottalized lexical items, 33 showed evidence of glottal squeak, and of these, only five exhibited glottal squeak as the only acoustic correlate of glottalization (less than 1% of all glottalized tokens).

For the nonprofessional Labnews speakers, it was not possible to determine whether any instances of glottal squeak occurred; a low level of noise was present throughout the recording, which made it impossible to determine whether this low-amplitude event had occurred. Only four of the remaining 10 speakers across both corpora showed squeak on any lexical items (Fig. 13); six speakers never produced this phenomenon. Three of the four nonprofessional speakers in the ABC corpus produced glottal squeak at least some of the time. Speaker SK produced it on over half of glottalized tokens (18 of 32 glottalized tokens, or 56%), a rate which was significantly greater than any other speaker (two-sided test; $p < 0.025$; $z \geq 2.310$). Speakers HH and RM showed rates of squeak on glottalized tokens of 29 and 8%, respectively, while ES never produced squeak. Among the professional speakers, five of the six radio news speakers did not produce squeak on any of the tokens examined in this study; speaker F3 produced squeak on 2 of 49 tokens.

The rates of employing different acoustic characteristics at different positions within the utterance (i.e. utterance-finally *vs.* utterance-medially) did not appear to vary systematically for our speakers.

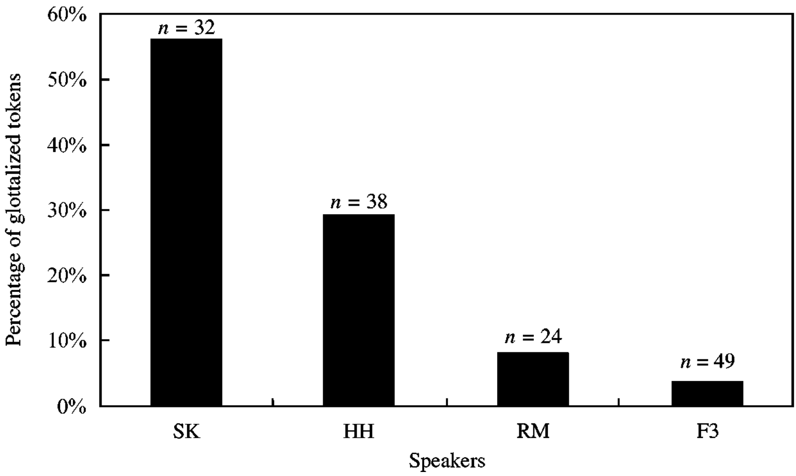

Figure 13. Percentage of glottalized tokens exhibiting glottal squeak, for four speakers. The remaining speakers did not produce glottal squeak on the lexical items examined.

3.4. *Rates of glottalization for intermediate vs. full intonational phrase-final position*

The rates of glottalization at the end of full intonational phrases (excluding utterance-final locations) were compared to the rates of glottalization at the end of intermediate intonational phrases for six of the 10 speakers. The rates of glottalization at full intonational phrase boundaries were higher than rates of glottalization at the ends of intermediate intonational phrases for all six speakers examined (Fig. 14). This difference was significant for four of the six speakers (one-sided test; $p < 0.001$, $z \geq 4.69$ for F2, FS, MM, and FJ).

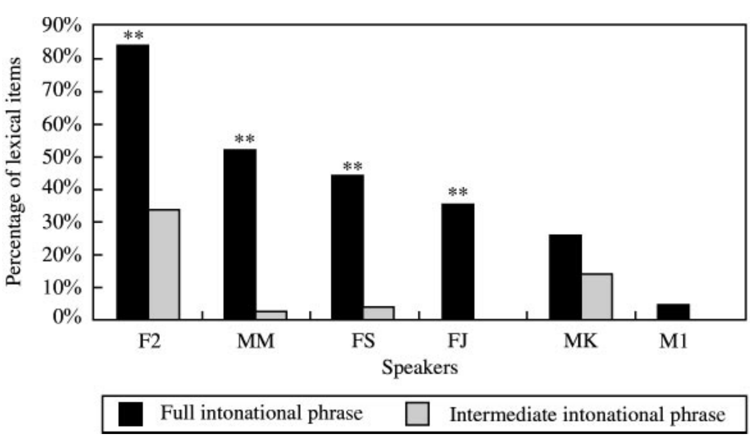

Figure 14. Rate of glottalization at utterance-medial full and intermediate intonational phrases in the Labnews corpus. ** = $p < 0.001$.

3.5. *Gender differences*

Results were mixed as to whether males or females produced more glottalization. Among the professional speakers, all three female speakers glottalized significantly more often than the three male speakers did (two-sided test; $p < 0.05$; $z \geq 2.107$). Among the nonprofessional speakers, results were more difficult to interpret. For the Labnews nonprofessional speakers, male speaker MM glottalized at a rate which was significantly higher than any other speaker, including male MK ($p < 0.005$, $z \geq 2.833$). Among the nonprofessional speakers in the ABC corpus, male speaker ES glottalized significantly more often than the female speakers ($p < 0.005$, $z \geq 3.032$).

4. Discussion

Results from this study extend earlier reports that normal speakers exhibit glottalized voice quality in association with the boundaries of intonation phrases of spoken American English utterances. Speech from a sample of professional and nonprofessional speakers showed a wide range of glottalization rates and significant differences in preferred acoustic characteristics. For example, some speakers produced glottal squeak and others did not; some produced diplophonia while others did not; and although speakers generally produced more aperiodic glottalizations than creak, some speakers showed a significantly greater asymmetry in this comparison than other speakers did.

Controlling for speaker identification, text and prosodic structure allowed us to examine the effects of other factors on glottalization near intonation phrase boundaries. For example, utterance-final intonation phrase boundaries were associated with higher glottalization rates than utterance-medial boundaries, and in utterance-medial position, full intonation phrases were glottalized more often than intermediate intonational phrases. The effects of other factors were more difficult to evaluate; evidence for the role of gender was inconclusive, as it was for the role of professionalism. Although there were hints that glottal squeak is dispreferred by professional radio news readers, and that female professional speakers may glottalize at higher rates than males, further work will be required to test these possibilities.

The extent of speaker variation in both rate and acoustic characteristics of glottalization raises several questions. For example, can we learn about individual variation in glottalization through comparison with earlier studies that have shown individual voice quality variation when using different elicitation tasks? How does this variation arise? And given the variation, how does phrase-level glottalization function in communication?

4.1. *Comparison with earlier studies of voice quality variation*

Our findings are consistent with earlier studies showing speaker variation in voice quality using other production tasks, in addition to the studies of glottalization discussed in the introduction section. For example, Stevens (1994) reported that speakers showed different patterns of change in the glottal waveforms for different degrees of prosodic prominence. Holmberg, Perkell, Hillman & Gress (1994) showed that speakers can produce the same SPL with varying glottal configurations and different levels of subglottal pressure. Gobl's (1988) study of emphatic stress showed that speakers differ in how they produce this type of prominence marking; spectra showed different characteristic degrees of spectral tilt at higher frequencies, and these contrasts have proven useful for synthesis (Karlsson, 1992).

The variability that we observed in the acoustic characteristics of regions perceived as glottalized is not limited to the tokens examined in this study. Rather, variability is a pervasive characteristic of glottalization phenomena, and was reported in all of the phonetically detailed studies we have surveyed. We surmise that speakers have some latitude in the way they choose to mark certain regions of an utterance by variations in voice quality. Considerable additional research will be necessary before we reach an understanding of precisely how, where, and why they make the choices they do. We think it is possible that similar mechanisms are responsible for speaker differences in voice quality in a variety of production tasks and measured parameters.

4.2. *Possible causes of variation in glottalization*

A potential cause of the variation observed here has to do with the articulatory mechanism itself, i.e., differences in the vocal apparatus which produces glottalization and other voice quality variations. One means by which creak or creaky voice may be achieved is by tightly approximating the arytenoid cartilages to allow vibration only over the anterior region of the vocal folds (Ladefoged, 1971; Stevens, 1977). This is often accompanied by an increase in high-frequency energy, and irregularity in timing between successive primary acoustic excitations may result (Scherer, 1994). Since individuals can vary in the detail of vocal anatomy (and perhaps in habits of neuromuscular control), physiological differences may contribute to a speaker's likelihood of producing phrase-level glottalization, and may influence preferences in its acoustic manifestation. Detailed models of vocal fold function, such as those developed by Titze & Talkin (1979) and by Hanson *et al.* (this volume), and Berry (this volume), will aid in understanding the variability in acoustic manifestations of glottalization gestures, and may clarify the effects of individual anatomical and physiological variation. It will be important to separate any contribution of individual speaker anatomy and physiology to variation in rate and acoustic characteristics of glottalization from the contributions of other linguistic factors such as structure and prominence, and socio-linguistic factors such as dialect and gender.

4.3. *Function of glottalization in communication*

One question that arises in connection with communicative function is the degree to which boundary-related glottalization is an independently planned event *vs.* a correlate of another planned event such as low subglottal pressure or low f_0, both of which are often observed at phrase boundaries. One piece of evidence that glottalization can be independently planned is that it is found in regions where f_0 is at the speaker's midrange or even rising. For example, Pierrehumbert & Talkin (1992) observed glottalization in word-initial vowels on a *H* H*-contour, Dilley *et al.* (1996) observed similar phenomena at *H** pitch accents, and the corpora described in this study included a number of tokens of boundary-related glottalization at *L–H%* boundaries, where the speaker's f_0 does not fall to the bottom of his or her range. However, the precise extent to which glottalization in normal speech is a planned event *vs.* a consequence of other factors remains to be determined.

Another question concerns the possible role of perceptual factors in shaping the range of individual variation in acoustic shape. Listeners appear to interpret a variety of acoustic evidence as suggestive of glottalized voice quality, including a drop in f_0 or amplitude, or the presence of irregular pitch periods (Hillenbrand & Houde, 1996; Pierrehumbert & Frisch, 1997; Rozsypal & Millar, 1979). If listeners accept a variety of acoustic cues as evidence of glottalization, this would allow speakers greater flexibility to produce glottalization by means of any number of (closely related) articulatory mechanisms which result in a variety of acoustic characteristics.

It is clear from these results and from earlier studies that phrase-level glottalization is not obligatory. However, it may enhance contrasts which are more directly signaled by other cues such as duration lengthening and boundary tones. If so, then glottalization may serve as a marker of "degree of finality" (when it occurs at phrase boundaries) or "degree of prominence" (when it occurs at pitch-accented syllables). Perceptual experiments will be necessary to evaluate the hypothesis that glottalization unrelated to segmental allophony is interpreted by listeners as evidence for a boundary or a prominence, and to determine whether it is interpreted along a continuum or as a contrastive binary feature.

Studies such as this one illustrate the importance of understanding individual speakers' phrase-level variations in voice quality in general, particularly glottalization. This understanding is especially significant for automatic speech recognition as well as f_0 estimation algorithms. Since glottalized regions often disrupt the periodicity on which these efforts largely rely, automatic detection of glottalized regions has been attempted. However, speaker variation makes this a difficult task. The observation by Kießling *et al.* (1993) that the glottalization detection algorithm they developed for one speaker did not generalize to a second speaker illustrates the importance of understanding the acoustic characteristics of an individual speaker's glottalizations. Better insight into variability in glottal events may also aid in the interpretation of phonetic evidence for voice-quality-based phonological contrasts (Vayra, 1994), and in understanding the relationship between normal and pathological pattern of nonmodal phonation. Finally, disentangling the contributions of individual speaker or gender-related differences in anatomy from those of socio-linguistic factors such as language, dialect and speaking style will be important in determining which aspects of phrase-level glottalization are part of the grammar, and which are paralinguistic in nature. These distinctions will impose significant constraints on cognitive models of both production and perception of speech. As is increasingly recognized, systematic variations in glottalization patterns will need to be accounted for in any comprehensive treatment of surface phonetic variation in normal speech.

References

Abramson, A. S. & Tingsabadh, K. (2000) Thai final stops: cross-language perception, *Phonetica*, **56,** 111–122.

Batliner, A., Burger, S., Johne, B. & Kießling, A. (1993) MÜSLI: a classification scheme for laryngealizations. *Proceedings of ESCA workshop on prosody*, Lund, pp. 176–179.

Beckman, M. & Ayers-Elam, G. (1997) Guidelines for ToBI Labeling. Unpublished manuscript, Ohio State University.

Beckman, M. E. & Pierrehumbert, J. B. (1996) Intonational structure in Japanese and English, *Phonology Yearbook*, **3,** 255–309.

Berry, D. A. (2001) Mechanisms of modal and nonmodal phonation, *Journal of Phonetics*, **29,** 431–450. doi:10.1006/jpho.2001.0148.

Byrd, D. (1994) Relations of sex and dialect to reduction, *Speech Communication*, **15,** 39–54.
Dilley, L. & Shattuck-Hufnagel, S. (1995) Variability in glottalization of word onset vowels in American English. In *Proceedings of the XIIIth international congress of phonetic sciences*, Stockholm, pp. 586–589.
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.
Docherty, G. J. & Foulkes, P. (1999) Derby and Newcastle: instrumental phonetics and variationist studies. In *Urban voices* (P. Foulkes & G. J. Docherty, editors), pp. 47–71. London: Arnold.
Docherty, G. J. & Foulkes, P. (1995) Acoustic profiling of glottal and glottalised variants of English stops. In *Proceedings of the XIIIth international congress of phonetic sciences*, Stockholm, pp. 350–353.
Fasold, R. W. (1981) The relation between black and white speech in the South, *American Speech*, **56**(3), 163–189.
Fischer-Jørgensen, E. (1989) Phonetic analysis of the stød in Standard Danish, *Phonetica*, **46,** 1–59.
Gerratt, B. & Kreiman, J. (2001) Toward a taxonomy of nonmodal phonation, *Journal of Phonetics*, **29,** 365–381. doi:10.1006/jpho.2001.0149.
Gimson, A. C. (1989) *An introduction to the pronunciation of English.* London: Edward Arnold.
Gobl, C. (1988) Voice source dynamics in connected speech, *Speech Transmission Laboratory Quarterly Progress and Status Report*, **1,** 123–159.
Gordon, M. & Ladefoged, P. (2001) Phonation types: a cross-linguistic overview, *Journal of Phonetics*, **29,** 383–406. doi:10.1006/jpho.2001.0147.
Hagen, A. (1997) *Linguistic functions of glottalizations and their language-specific use in English and German.* Master's thesis, Universität Erlangen-Nürnberg.
Hanson, H. M., Stevens, K. N., Kuo, H.-K. J., Chen, M. Y. & Slifka, J. (2001) Towards models of phonation, *Journal of Phonetics*, **29,** 451–480. doi:10.1006/jpho.2001.0146.
Henton, C. G. & Bladon, A. (1987) Creak as a sociophonetic marker. In *Language, speech and mind: studies in honour of Victoria A. Fromkin* (L. Hyman & C. Li, editors), pp. 3–29. London: Routledge.
Hillenbrand, J. M. & Houde, R. A. (1996) The role of f_0 and amplitude in the perception of intervocalic glottal stops, *Journal of Speech and Hearing Research*, **39,** 1182–1190.
Hollien, H. (1974) On vocal registers, *Journal of Phonetics*, **2,** 125–143.
Holmberg, E. B., Perkell, J. S., Hillman, R. E. & Gress, C. (1994) Individual variation in measures of voice, *Phonetica*, **51,** 30–37.
Huber, D. (1988) *Aspects of the communicative function of voice in text intonation.* PhD thesis, Chalmers University, Göteborg/Lund.
Karlsson, I. (1992) Modelling voice variations in female speech synthesis, *Speech Communication*, **11,** 491–495.
Kerswill, P. E. (1987) Levels of linguistic variation in Durham, *Journal of Linguistics*, **23,** 25–49.
Kießling, A., Kompe, R., Niemann, H., Nöth, E. & Batliner, A. (1993) Voice source state as a source of information in speech recognition: detection of laryngealizations. In *Proceedings of NATO ASI*, Budion.
Kirk, P. L., Ladefoged, P. & Ladefoged, J. (1984) Using a spectrograph for measures of phonation types in a natural language, *UCLA Working Papers in Phonetics*, **59,** 102–113.
Klatt, D. H. & Klatt L. C. (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers, *Journal of the Acoustical Society of America*, **87,** 820–857.
Kohl, A. & Anderson, B. (2000) Glottalization as a sociolinguistic variable in Detroit. *Proceedings of NWAV*, 29, Michigan State University.
Kohler, K. J. (1994) Glottal stops and glottalization in German, *Phonetica*, **51,** 38–51.
Kohler, K. J. (2000) Investigating unscripted speech: implications for phonetics and phonology, *Phonetica*, **57,** 85–94.
Ladefoged, P. (1971) *Preliminaries to Linguistic Phonetics.* Chicago: University of Chicago Press.
Ladefoged, P. & Maddieson, I. (1996) *The Sounds of the World's Languages.* Oxford: Blackwell.
Laver, J. (1980) *The Phonetic Description of Voice Quality.* Cambridge: Cambridge University Press.
Local, J. K., Kelly, J. & Wells, W. H. G. (1986) Towards a phonology of conversation: turn-taking in Tyneside English, *Journal of Linguistics*, **22,** 411–437.
Local, J. K., Wells, W. H. G. & Sebba, M. (1985) Phonology for conversation: phonetic aspects of turn delimitation in London Jamaican, *Journal of Pragmatics*, **9,** 309–330.
Pierrehumbert, J. (1995) Prosodic effects on glottal allophones. In *Vocal fold physiology: voice quality control* (O. Fujimura & M. Hirano, editors), pp. 39–60. San Diego: Singular Publishing Group.
Pierrehumbert, J. & Frisch, S. (1997) Synthesizing allophonic glottalization. In *Progress in speech synthesis* (J. H. van Santen, R. W. Sproat, J. Olive & J. Hirschberg, editors), pp. 9–26. New York: Springer.
Pierrehumbert, J. & Talkin, D. (1992) Lenition of /h/ and glottal stop. In *Papers in laboratory phonology II* (G. Docherty & D. R. Ladd, editors), pp. 90–117. Cambridge: Cambridge University Press.
Pollack, I. (1971) Amplitude and time jitter thresholds for rectangular wave trains, *Journal of the Acoustical Society of America*, **50,** 1133–1142.
Rozsypal, A. J. & Millar, B. F. (1979) Perception of jitter and shimmer in synthetic vowels, *Journal of Phonetics*, **7,** 343–355.
Scherer, R. (1994) Laryngeal function during phonation, *NCVS Status and Progress Report*, **7,** 171–186.

Shattuck-Hufnagel, S. & Turk, A. (1998) The domain of phrase-final lengthening in English. In *Proceedings of the 135th meeting of the Acoustical Society of America*, Seattle, Washington, Vol. 2, pp. 1235–1236.

Stevens, K. N. (1977) Physics of laryngeal behavior and larynx modes, *Phonetica*, **34,** 264–279.

Stevens, K. N. (1994) Prosodic influences on glottal waveform: preliminary data. In *Proceedings of the international symposium on prosody*, Yokohama, Japan, pp. 53–64.

Titze, I. R. & Talkin, D. (1979) A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation, *Journal of the Acoustical Society of America*, **66,** 60–74.

Turk, A. (1999) Structural influences on boundary-related lengthening in English. In *Proceedings of the XIVth international congress of phonetic sciences*, San Francisco, Vol. 1, pp. 237–240.

Umeda, N. (1978) Occurrence of glottal stops in fluent speech, *Journal of the Acoustical Society of America*, **64**(1), 88–94.

Vayra, M. (1994) Phonetic explanations in phonology: laryngealization as the case for glottal stops in Italian word-final stressed syllables. In *Phonological 1992. Proceedings of the 7th international phonology meeting*, Krems, Austria (W. Dressler, M. Prinzhorn & J. Rennison, editors), pp. 275–293. Torino: Rosenberg & Sellier.

Zue, V. W. & Seneff, S. (1988) Transcription and alignment of the TIMIT database. In *Proceedings of the second meeting on advanced man–machine interface through spoken language*, Hawaii, U.S.A. pp. 11.1–11.10.