

# The Effect of Glottalization on Voice Preference

*Hongwei Ding, Oliver Jokisch and Rüdiger Hoffmann*

Laboratory of Acoustics and Speech Communication  
Dresden University of Technology

{hongwei.ding; oliver.jokisch; ruediger.hoffmann}@ias.et.tu-dresden.de

## Abstract

The impact of phrasal prosody on glottalization is documented in many publications ([4], [7]). Besides prosodic boundary and stress, other influencing factors such as the speaking style have been studied in [8]. The work reported here examines the relationship between the objective preference of listeners and the occurrence of speakers' glottalization, speech data employed for this purpose were phonetically balanced sentences in six languages. Additional experiment, concerning the influence of reading style on glottalization, was conducted with American words and phrases read with monotone and isolated Chinese syllables segmented from carrier sentences. Evaluating the statistics from this investigation, we can come to following conclusions: (a) The occurrence and degree of glottalization can be different across speakers. (b) As an prosodic effect, glottalization is not undesired for speakers. (c) A well-defined reading style can increase the occurrence.

## 1. Introduction

The term glottalization used in different studies has different acoustic characters. In this paper, glottalization is understood to be a region in the speech signal with irregular pitch periods, and often accompanied by extremely low  $f_0$ , voicelessness or pause.

The fact that speakers exhibit irregular speech signal, which is referred to as glottalization, in normal speech has been observed in many languages: in English [2], German [4], Chinese [3] and many other languages [6]. Glottalization can serve as a communicative function in natural speech, if it can be reproduced in speech synthesis, the naturalness of synthesis can also be improved.

In dealing with the databases of speaker selection for multilingual speech synthesis [3], we came across the problem of glottalization. We wondered whether the occurrences of glottalization destroy the preference of voice for the listeners, and whether the occurrence and degree of glottalization of an individual speaker can be influenced by different reading styles.

## 2. Investigation method

In our experiments, native speakers were invited to our recording studio to read several phonetic balanced sentences. Most of them were professional speakers. None of them spoke with a habitual glottalized voice.

### 2.1. Speech database

The speech material can be divided into two parts, the first part was employed for the study of the relationship between glottalization occurrences and the preferred voice of listeners.

- 6 sentences of 7 speakers of American English
- 6 sentences of 6 speakers of British English
- 5 sentences of 8 speakers of Spanish
- 5 sentences of 5 speakers of Italian
- 5 sentences of 5 speakers of Dutch
- 12 sentences of 7 speakers of Chinese

The second part was devoted to the study of the relationship between glottalization and reading style from 2 speakers.

- 2008 words and phrases uttered by the selected American speaker
- 1644 syllables segmented from carrier sentences uttered by the selected Chinese speaker

### 2.2. Preference Tests

The above-mentioned first part of the databases was used in the listening test. The number of listeners was different for each language:

- 24 native and 22 non-native listeners (American English)
- 20 native and 5 non-native listeners (British English)
- 20 native and 13 non-native listeners (Spanish)
- 20 native and 16 non-native listeners (Italian)
- 20 native and 20 non-native listeners (Dutch)
- 20 native and 10 non-native listeners (Chinese)

The sentences were presented in A/B pair comparison (same utterance spoken by two different speakers) and in random order. All speaker combinations have been tested for the mentioned test sentences. The listeners used headphones and made their judgements on PC keyboard. They were only asked for their overall impression/decision and each decision for a speaker's example was scored by one point. Neutral decisions were disabled. The overall ranking of the speakers was derived from the sum of the speaker's decision points. The speaker with the highest score was selected as preferred voice. Actually, the listeners judged the speakers intuitively according to their prosody, reading style and the pronunciation as a whole.

### 2.3. Classification of glottalized events

In order to conduct the investigation, it is important to know the circumstances where glottalization occurs. According to the findings in different languages [2], [3], [4], glottalized voice quality can appear in a variety of locations, it often occurs at the vowel initial, especially V-V hiatus, plosives and at the end of phrases or utterances. In order to capture different occurrences of glottalization, Kohler [4] divided it into three categories in terms of prosody:

- Plosive prosody: glottalization of plosives;
- Word prosody: glottalization of word-initial vowels;
- Utterance prosody: glottalization of phrasal finals.

Our labelling experiences also confirmed, that these are circumstances for the occurrences of glottalization, and the categorization is also favorable to account for the effects of prosody. But in this work our interest is focused on glottalization of vowels, that of consonants would be neglected. Because of the special characteristics of Chinese, a lexical tone 3 prosody is added. The glottalization of tone 3 occurs in the middle, the whole vowel part can be affected by the irregularity. The onset and offset glottalization of tone 2 and 4 was neglected. Therefore we have two categories for non-tonal languages, three categories for Chinese, which is described in Table 1.

Table 1: *Locations of glottalization investigated.*

Location	Constraints
Lexical tone 3 prosody	glottalization of tone 3 (only in the case of Chinese)
Word prosody	glottalization of word-initial vowels
Utterance prosody	phrasal final glottalization (often with low f0)

Sometimes it is very difficult to decide into which category a glottalization should fall. Because there are many third tones, which have no glottalization if they are not phrasal finals. It seems utterance prosody has stronger influences on the syllable than lexical tones. Therefore in our statistics, if a glottalized tone 3 was also a phrasal final, it was counted as utterance prosody. Every occurrence of glottalization was counted only once. The category was carefully considered by experts.

Although there exist some automatic glottalization labelling algorithms, for example in [1], the accuracy still requires retesting. In order to acquire the most accurate information, all the material was carefully manually labelled by expert. Word fragments were acoustically analyzed in terms of presence/absence of glottalization, using spectrographic analysis, the F0 value and listening confirmation as the same time.

### 3. Results

The results of the investigation are presented separately in two sections: the comparison of speakers and the glottalization in special reading styles.

#### 3.1. Comparison of Speakers

The frequencies of the glottalization of speakers in each language will be demonstrated in a diagram. All the frequencies of glottalization for European languages are divided into two categories, with the upper column indicating the frequencies of vowel-initial glottalization and the lower column showing phrasal final glottalization.

##### 3.1.1. American English and British English

The comparison of glottalization occurrences of American speakers can be observed in Fig. 1. The most favorite speaker was speaker 3. She was a linguist and also a professional speaker for theater. She could read the sentences in a clear way.

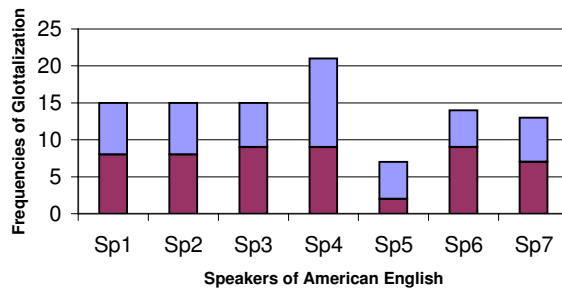


Figure 1: *Frequencies of glottalization of American Speakers. (Upper column: vowel initial, Lower column: phrase final)*

She had the second place in the glottalization occurrences. She employed glottalized voice between V-V hiatus as well as at phrasal finals. Here is an example of glottalized segment at the end of an utterance. In the sentence "One of the strangers is in danger.", a clear glottalization can be observed in word "danger" in Fig. 2.

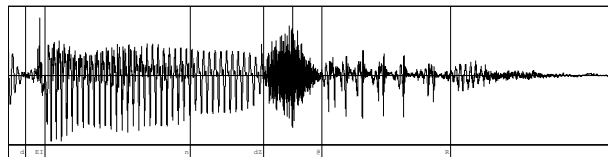


Figure 2: *Waveform of "danger" of US-speaker 3.*

Whereas a regular speech signal with declining amplitude can be observed in the correspondent word fragment of speaker 5 in Fig. 3.

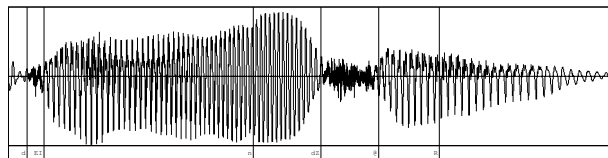


Figure 3: *Waveform of "danger" of US-speaker 5.*

The comparison of the British speakers can be shown in Fig. 4. The most favorite speaker was speaker 6, who was a speaker for television. She had 14 occurrences of glottalization, had also the second place in occurrences. Glottalization of V-V hiatus in "(cen)ter of" /t@Q@v/ from the sentence: "Look through the center of the moon." can be observed in Fig. 5.

But another speaker had a smooth V-V transition in the correspondent phrase in Fig. 6.

##### 3.1.2. Spanish Italian Dutch

As to the other European languages, the comparison of speakers can be observed in following diagrams: Spanish in Fig. 7, Italian in Fig. 8 and Dutch in Fig. 9.

The most favorite speakers were speaker 4 in Spanish, speaker 4 in Italian and speaker 1 in Dutch. All of them were among the speakers who provided the most glottalization.

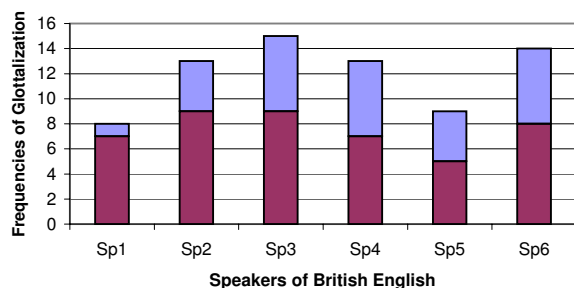


Figure 4: Frequencies of glottalization of British Speakers. (Upper column: vowel initial, Lower column: phrase final)

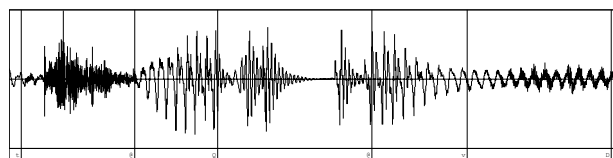


Figure 5: Waveform of "(cen)ter of" of UK-speaker 6.

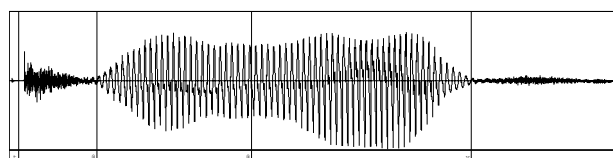


Figure 6: Waveform of "(cen)ter of" of UK-speaker 1.

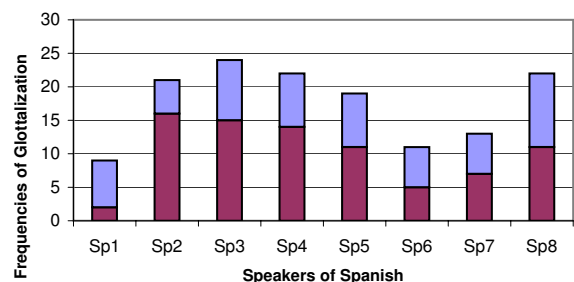


Figure 7: Frequencies of glottalization of Spanish Speakers. (Upper column: vowel initial, Lower column: phrase final)

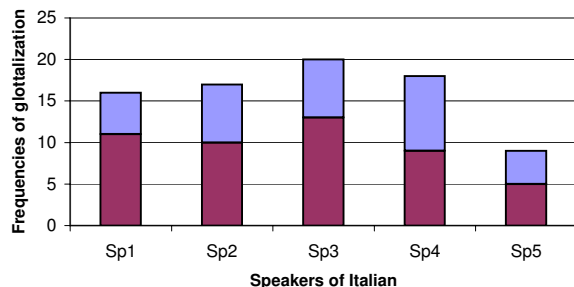


Figure 8: Frequencies of glottalization of Italian Speakers. (Upper column: vowel initial, Lower column: phrase final)

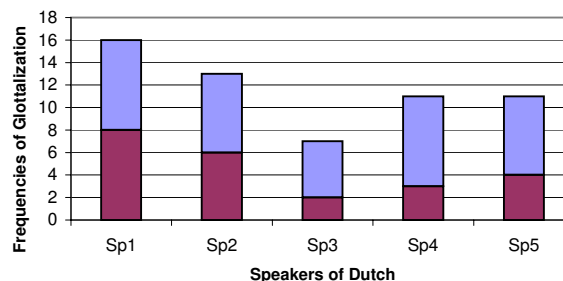


Figure 9: Frequencies of glottalization of Dutch Speakers. (Upper column: vowel initial, Lower column: phrase final)

### 3.1.3. Chinese

The same phenomenon can also be observed in Chinese in Fig. 10 with the upper column indicating glottalization of tone 3, middle column of vowel-initial and the lower column of phrasal final. If other tones demonstrated glottalized gestures at the end of phrases, they would also be counted as utterance final glottalization as long as the glottalization occurs not only as offset.

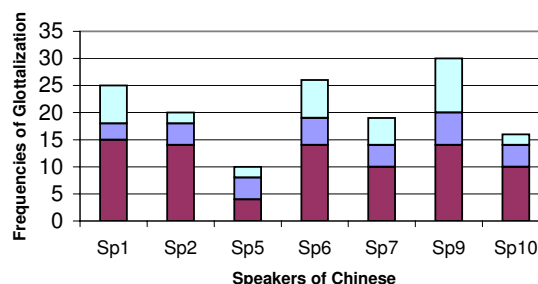


Figure 10: Frequencies of glottalization of Chinese Speakers. (Upper: tone 3, Middle: vowel-initial, Lower: phrase final)

The most favorite speaker was speaker 1, who had the third place in the frequency of glottalization. She was also a professional television speaker, she usually employed glottalization when  $f_0$  dips down, in the middle of tone 3, at the end of tone 4 and at the end of phrase. Here is an example of word "sheng1chu4(deep site)" at the phrasal final in Fig. 11.

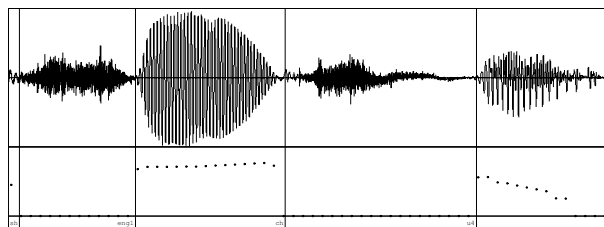


Figure 11: Waveform and  $f_0$  of "sheng1chu4" of speaker 1.

The correspondent word segment from speaker 5, who has little glottalization, showed a smooth tone 4 contour with  $f_0$  dipping down and amplitude gradually declining in Fig. 12.

## 3.2. Characteristics in special reading styles

The most favorite speakers of European languages were then asked to read words and phrases with monotone. But in this in-

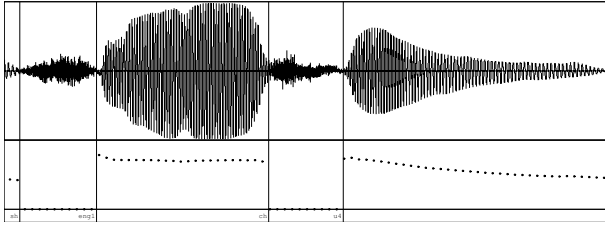


Figure 12: Waveform and  $f_0$  of "sheng1chu4" of speaker 5.

vestigation, only the American and Chinese speech data were studied, because the data in the other European languages demonstrated the similar glottalization effect as those of American English.

### 3.2.1. American Speaker

The selected speaker was asked to read 2008 words and phrases in a clear and monotonous way. These speech phrases were observed with following characteristics:

- The absence of phrasal final glottalization was striking, because of lacking of phrasal prosody.
- On the other hand, V-V hiatus glottalization were enhanced due to the clear reading style.
- The degree of vowel initial glottalization is greatly decreased.

### 3.2.2. Chinese Speaker

The selected Chinese speaker was then asked to read 1644 carrier sentences with the investigated syllables (represented as \*) embedded in "Zhe4ge5 \*zi4 zhen1nan2nian4" (This character \* is really difficult to read.). She read in such a way that the investigated syllable was clearly articulated. All 411 third tone syllables had more or less glottalization in the middle of the syllable when  $f_0$  goes down. But they demonstrated different acoustic characteristics. For example: a typical glottalized tone 3, just like tone 3 at phrasal finals, can be observed in Fig. 13; glottalization with voiceless parts can be observed in most syllables with /a/ in Fig. 14,

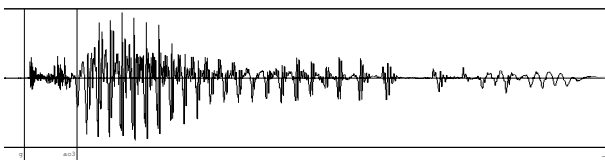


Figure 13: Waveform of "gao3".

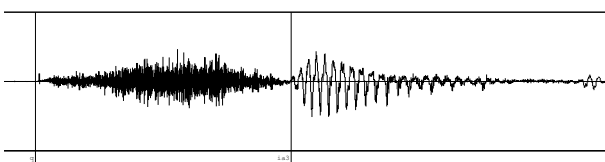


Figure 14: Waveform of "qia3".

## 4. Conclusion

Comparing the voice preference of the listeners and the glottalization occurrences of the speakers, we found the most preferred speakers of listeners also employed frequent glottalization.

This paper provides a cross-language study to confirm and extend previous researchers' observations:

- The occurrence and degree of glottalization are different across speakers [2] [3] [4].
- V-V hiatus glottalization can be enhanced by a well-defined reading style [8].

On the other hand, this work claims to enrich the understanding of glottalization by investigating glottalization in a multilingual voice selection.

- The employment of moderate glottalization to express a natural prosody was quite accepted by listeners.
- In tone languages, besides phrasal prosody, the lexical prosody is also responsible for the occurrences of glottalization. The most frequent occurrences of glottalization in Chinese is related to an extreme low  $f_0$ , when the speaker dips down her fundamental frequency to express the lexical prosody as well as utterance prosody, because at the end of utterance, not only tone 3 but also other tones can also be glottalized.

In general the paper should provide a better understanding of glottalization, which will contribute to the treatment of surface phonetic variation in speech, the selection of speaker and the generation of acoustic inventories. Further studies will be continued to investigate the modelling of irregular speech periods/epochs with regard to a concatenative speech synthesis.

## 5. References

- [1] BATLINER, A. & BURGER, S.& JOHNE, B.& KIESSLING, A. (1993): MÜSLI: A Classification Scheme For Laryngealizations. *ESCS Workshop on Prosody*. Lund, 1993. 176-179
- [2] DILLEY, L. & SHATTUCK-HUFNAGEL, S. & OSTENDORF, M. (1996): Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*. Vol. 24, 423-444
- [3] DING, H. & JOKISCH, O. & HOFFMANN, R. (2000): Glottalization in inventory construction: A cross-language study. *ISCSLP*. Hongkong, 2000. 37-40
- [4] KOHLER, K.J. (1996): Glottalizations across Languages. *Sound Patterns in Spontaneous Speech, Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK)* 30, 159-193
- [5] KOHLER, K.J. & PETERS, B. & WESENER, T. (2001): "Interruption glottalization in German spontaneous speech", In *DISS'01*. 45-48. Disfluency in Spontaneous Speech August 29-31, 2001 Edinburgh, Scotland, UK
- [6] LADEFOGED, P. & MADDIESON, I. (1996): *The Sounds of the World's Language*. Blackwell Publishers
- [7] PIERREHUMBERT & TALKIN (1992): Lenition of /h/ and Glottal Stop. In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. ed. by G. Doherty and D. R. Ladd, 90 - 117. Cambridge: Cambridge University Press
- [8] PIERREHUMBERT, J. (2000): #, +, and ? . In *7th meeting on Laboratory Phonology*. Nijmegen. June, 2000