



The Influence of language and speech task upon creaky voice use among six young American women learning French

Benoist-lucy Agathe¹ Pillot-loiseau Claire¹

¹ Laboratoire de Phonétique et Phonologie, UMR 7018, CNRS/ Univ Sorbonne Nouvelle-Paris 3

agathebenoistlucy@yahoo.fr, claire.pillot@univ-paris3.fr

Abstract

This article intends to study the influence of two parameters (language and speech task) on the manifestation of creaky voice among six American women learning French. The speakers were recorded in both reading and spontaneous conditions, in English and French. These languages were chosen for their difference at the phonatory level. The number of creaky voice occurrences, and creakiness timing were calculated. The aim was to answer the following questions: Can American native speakers learning French show a differentiated use of creaky voice depending on the language they speak? Can American native speakers show a variable use of creaky voice whether they read a text or talk spontaneously? Our study shows that the six American speakers of this research show significantly more creaky voice in English than in French and in spontaneous speech than in the reading task. This leads us to consider that phonatory aspects of a language are part of a language itself and that certain speech conditions are more prone to exhibit extreme vocal effects such as creaky voice, given their propensity to expressiveness.

Index Terms: creaky voice, L2, American English, French, learners

1. Introduction

Creaky voice is described by Hollien and Michel [1] as a «*train of discrete laryngeal excitations, or pulses of low frequency*». This phonation type is physiologically characterized by an irregular glottal cycle and a longer closed phase than in modal voice [2] [3] [4]. Vocal folds are thicker and shorter and vibrate at a lower frequency. Perceptively, it conveys an impression of sudden decrease of the pitch [5].

Creaky voice has been considered under various angles over the years. On a pathological level, it was viewed as the manifestation of voice alteration caused by granuloma [6]. On a phonological level, it was studied as a contrast [7]. On a conversational level, it was viewed as a demarcative function [8] [9] [10], and a turn taking cue [11]. On a pragmatic level, it was perceived as the expression of specific emotions (surprise, admiration or suffering [12]) or as a mark of hesitation [13]. It was also lately presented as a social function [14] [15] [16]. Indeed, according to Esling [17], the comparison of voice quality aspects makes a regional, social or linguistic group recognizable. In this context, voice quality could be considered as an identifier of a group, and dynamic voice aspects, such as intonation or phonation effects for example, can thus convey paralinguistic information. For these specific reasons, the creaky voice phenomenon doesn't appear in all languages with the same proportion. American English is one of the languages for which creaky voice is described [18] [5].

In this study, our interest thus concerns American English but also French, given the specificity of our population.

2. State of the art

Studies relating the creaky voice phenomenon to social and paralinguistic features started with Edward Sapir, who brought the idea that vocal specificities are a «*social scheme*» reflecting the accent of an entire group: «*There is always something about the voice that must be ascribed to the social background*» [19]. He was then followed by Esling [17], Knowles [20] and Stuart Smith [21] in that same line of research. The creaky voice conception evolved over the years, passing from a pathological phonation type to an accepted aspect of healthy speech [16] and from an elected masculine specificity to a possible feminine setting [22] [23] [24] [15] [16]. As for the recent studies on American English, creaky voice could even be, according to Yuasa [15], a sociocultural habit common among young American women thus introducing a «*new voice quality for young urban oriented upwardly mobile American women*». The conclusion of this research was that female speakers of American English residing in California employed creaky voice much more frequently than comparable American male and Japanese female speakers. Other studies also underline the frequent use of this phonatory register. This creaky voice phenomenon is indeed so prevalent that it could, according to Böhme and Shattuck-Hufnagel, be a solid cue to speaker's identification [25]. Articles also report its use as a stylistic device [16] [24] [14]. As for French, phonatory settings are described as nasalized, breathy and sometimes whispery [26] but apart from hesitation manifestations, there is no mention of a possible creaky voice phenomenon in this language.

3. Objectives

The aim of this article is to take into account the difference between French and English on the vocal quality level, in order to study the effect of language on the creaky voice manifestation. As we said, American speakers are more prone to turn to creaky voice, which is now considered by some as an important part of self-expression. In the light of social schemes induced by this language, we can reasonably expect to find a differentiated use of creaky voice depending on the language used by a same speaker. We suppose that, given the phonatory habits of our American native speakers, and the scarcity of the creaky voice phenomenon effect in French, American speakers learning French as a foreign language might show a probable imbalance in their creaky voice use. We thus wondered whether the creaky voice phenomenon could vary depending on the language spoken. Is it American English specific or can it be observed when American learners speak French? Do they turn to creaky voice with the same frequency in the two languages? The hypothesis supposes that L2 learners take a language's phonation type schemes into

account as they speak. We thus expect to find more creaky voice when the speakers talk English than French.

A second specificity will be studied in this article. The great majority of studies led on creaky voice, only focused on read speech and mentioned the importance of testing the manifestation of creaky voice in a spontaneous speech mode [16]. By creating a corpus with both a reading task and a spontaneous speech task, we wanted to confront the manifestation of creaky voice between a reading and a conversational situation. We wanted to investigate whether the creaky voice phenomenon could vary depending on the speech task. According to Wolk *et al.* [16], less creaky voice is expected in the reading task, in comparison to spontaneous speech. The hypothesis is that the latter condition would provide more occurrences of creaky voice because of its propensity to develop expressional reactions. This comparison would be a way to confirm or rule out Wolk's hypothesis, according to which more creaky voice is expected in spontaneous speech than in the reading task because the person is more expressive [14]. It would also be a way to underline a distinction between a social use of creaky voice against a linguistic one (creaky voice as a boundary marker). In sum, we are wondering whether the appearance or proportion of creaky voice can vary depending of the language spoken or the speech modality.

4. Method

We recruited six. American females who were part of the same exchange program in Paris. They were all 20 years old. This age condition corresponded to the literature consensus stating that the creaky voice effect is more robust among American young women and often occurs in a College context [15]. They had arrived in France five months before the recording sessions took place, and they were all living in a French family. They had all studied French since High school. Their foreign language level was estimated according to the Common European Framework of Reference for Languages as "independent users" (B1 or B2).

The recording session took place in a sound proof room. The speakers were asked to attend in the middle of the day in order to avoid vocal harshness that can result from morning phonation. In order to create a favorable situation for conversational speech, we asked them to come by pairs and to choose the person they wanted to be recorded with. The complicity of the speakers was important in the frame of the study so we made sure these conditions were respected.



Figure 1 Photography of two speakers during the recording session

We used the Pro Tools cross-platform workstation which allowed simultaneous recording of two audio acoustic tracks and two electro-glottographic tracks. Given the difficulty to acoustically process creaky voice, we considered the EGG as a way to convey accurate information about the glottal cycle. The microphones were C520L helmet microphones (AKG brand) which allowed the mouth to microphone distance to remain stable. The sampling frequency was 44100 Hz. The electro-glottographic graphs were Glottal Enterprises EG2-PCX2. The data was gathered and saved on an iMacOSX (Intel Core i5, 4Go).

The corpus was composed of the reading of the text « The North Wind and the Sun » read in French and English (repeated three times in each condition). We also collected about fifteen minutes of spontaneous speech in each language. Given the speakers' complicity we didn't have to induce any conversation. We only asked them to switch to the other language so that the time in each language was respected. At the end of the recording session, we asked our speakers to fill out a questionnaire, which collected information about the linguistic profile of the speakers. None of the six speakers knew about the nature of the research.

The data was then annotated on the Praat software: we extracted the creaky passages and talking time for each speaker. This was performed by one single auditor who used a headphone set (Sony MDR-V150) in order to guarantee the perceptive quality. We annotated different levels: the creak portion on the first tier, the phonemic context on the second, the orthographic phrase on the third and the speaker's name on the fourth as shown below.

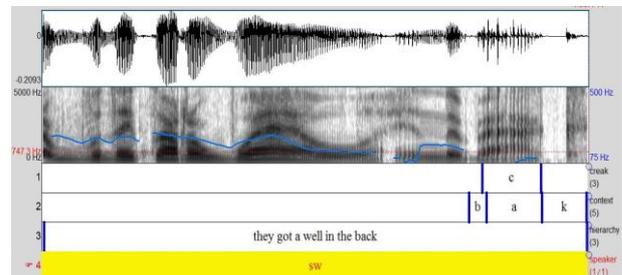


Figure 2 Annotation for the spontaneous sentence "They got a well in the back" (speaker S.W.)

For the accounting of the creaky passages, we used a recommended method in the literature [27], that is to say a perceptive method led by the confirmation through the visualization of acoustic cues. The idea is to listen to the recordings and select all the passages corresponding to the auditory criterion of a "rough quality with the additional sensation of repeating impulses" [12]. In a second time, we certify this auditory aspect through the visualization of spectrograms and F0 contours by observing long glottal cycles, fundamental frequency decrease or disappearance and wave sound specific shape. We finally turned to EGG waveform to ultimately confirm these creaky passages and segment them. The accounting of creaky voice occurrences is a delicate notion because the reference system varies from one author to another. We considered an occurrence any creak extract uninterrupted by modal phonation.

5. Results

The annotation revealed **960** occurrences of creaky voice for the six speakers pooled together. These occurrences showed this distribution:

	English	French
Reading task	126	68
Conversational task	574	192

Table 1 Occurrence number of creaky voice portions according to language and reading task

It is not conceivable to compare such rough numbers, especially between reading and conversational tasks, which don't exhibit the same speaking length.

The methodology for creaky voice accounting varies a lot from one author to another. Some use the number of creaky syllables on the number of total **syllables** [8]; other authors compare the number of **utterances** showing creaky voice phonation to the total number of utterances [14]; other count the number of **words** containing creakiness and calculate a ratio with the total number of words [15]. It is a delicate point because creaky voice doesn't necessarily stick to these syntactic boundaries.

For that specific reason, we decided to follow Kane and Drugman's method [27], consisting in calculating the ratio of cumulated creak time (all the cumulated creaky portions) divided by the total spoken time. As we said, we considered an occurrence any creak extract uninterrupted by modal phonation. We calculated the cumulated creak duration and then divided it by the total duration for each speaker. We thus obtained a ratio indicating the amount of creak in the speaker's speech for each speaker.

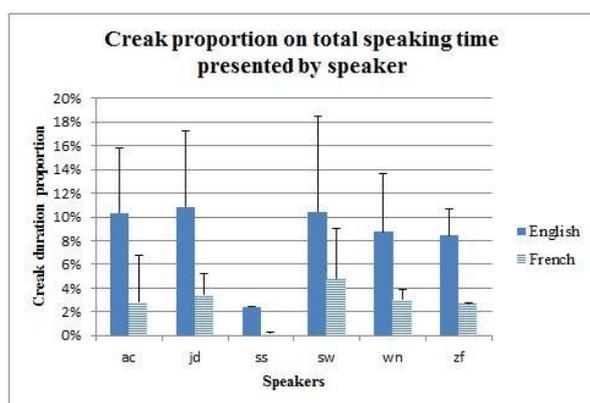


Figure 3 Creak duration on total speaking time by speaker

The first observation we can make is an important inter-speaker variability: the creak proportion can indeed go from 0,2 to 10% depending on the speaker and language concerned.

We then calculated these ratios in each speech task and in each language.

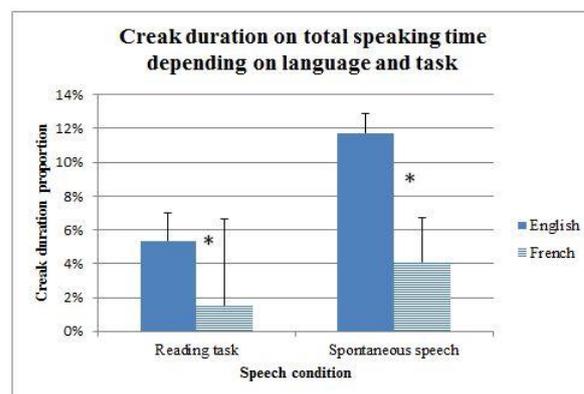


Figure 4 Creak duration on total speaking time (for language and task)

We used a two way ANOVA testing the effect of language and modality on the creak proportion. The significance for both language ($F(1,20)=21,32$ with $p<0,002$) and speech task ($F(1,20)=12,91$ with $p<0,018$) highlighted a significantly higher rate of creakiness in English for the language variable, and a higher rate of creakiness for spontaneous speech for the speech task condition. Clearly, it means that the speakers tend to display more creakiness when they speak English than French, and when they are in a conversational speech condition compared to when they read. We notice that the condition in which the creaky rate is the highest (11,7%) is when the speakers speak English and are in a spontaneous speech condition whereas the lowest rate concerns French during a reading condition.

This ratio is convenient for giving us an idea of the proportion of creakiness in terms of duration, but deprives us from having an idea of the number of occurrences. We thus turned to another accounting method that allowed this representation: the proportion of occurrences by calculating the number of occurrences divided by the total speaking length. In this method, the higher the ratio is, the more occurrences there is.

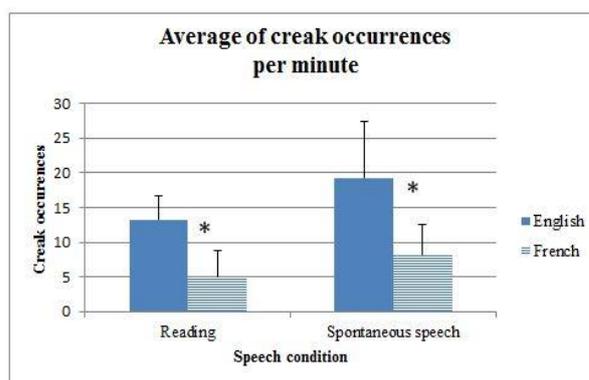


Figure 5 Average of creak occurrences (per minute)

We also used a two ways ANOVA testing the effect of language and modality on the creak occurrence proportion. The significance for both languages ($F(1,20)= 19,84$ with $p<0,002$) and speech task ($F(1,20)= 4,505$ with $p<0,046$) illustrated the same trend than for the previous analysis,

highlighting a significant higher rate of creakiness for English, and a higher rate of creakiness for spontaneous speech.

6. Discussion

The presence of creaky voice among American speakers has been depicted in the literature over the years. By choosing six American English speakers, we expected to find occurrences of this phonation type in our recordings. However, what we did not know was the phonatory attitude the speakers would display when switching to another non-native language. The hypothesis was that they would show less creaky voice in French because this phonatory effect is described as part of the language setting in English, but absent in French.

The results gathered here underline a discrepancy between the two languages confirming this idea. It is difficult to compare these results to those of previous studies because no studies (that we know of) have compared two languages for a same speaker on the basis of creaky voice. Yuasa [15], who compared American subjects to Japanese ones, found an important difference between the English American speakers and their Japanese counterparts. Her results showed that young American women used creaky voice twice as much as Japanese women. Given the fact that her speakers were not the same person in two different languages, her results are exposed to inter-individual variations. By using foreign language learners, we offered the possibility to keep this inter-individuality variation source controlled and observe a strictly language-induced difference. These results urge us to consider language according to the « vocal quality » notion developed by Laver [28], who depicted the importance, for the production of a language, of not only the supra-laryngeal settings (modulation of the source), but of the laryngeal settings (source) as well. This is especially true for learners of a foreign language, who should stick to a language's requisites at both levels. As Honikman described it [26], the difference between French and English on the phonation aspect leads to the idea that phonatory effects are language dependent and that they have to be taken into account when one speaks a language. Our results encourage us to consider that the speakers integrated the creaky phonatory aspect in their speech when they spoke English, and reduced it in French because they noticed its non-relevancy in this language. We can also consider the possibility that this inter-language difference might be induced by the cognitive weight of speaking a foreign language, which might possibly leave more vigilance for specific phonatory use.

The other aspect that we wanted to examine in this article was the possible effect of a speech task on the creaky voice manifestation. The hypothesis raised in the literature was that more creaky voice was expected in spontaneous speech because it was more likely to exhibit expressivity aspects [15]. No study compared the two speech conditions for same speakers but Podesva [14] compared the use of phonatory registers (creak and falsetto) for an American speaker in two situations of different formality (the first one at a barbecue party with friends, the second one during a medical consultation). The results of the study show that the manifestation of falsetto and creaky voice varies significantly depending on the formality of situation, showing 9,07% for the informal and 3,72% for the formal situation. According to the author, phonatory registers such as creak and falsetto can be a

resource to manage our position in the social environment. Even though the formality of situation cannot be compared to a speech task, we cannot help to underline, in our results, a discrepancy in the use of creaky voice between the reading task and the spontaneous conversation. This difference could thus be attributed to the fact that a situation which favors direct interaction (and thus, expressivity) also encourages use of more extreme pitches which can influence the appearance of creaky voice.

These results cannot be analyzed without bearing in mind variations of creaky voice use. This is especially true for a vast country like the United States, which shows differentiated recourse to this phonatory aspect depending on states and regions [16]. Apart from the regionally or geographically induced variation, some authors also underlined very strong individual variations [29] [2]. This effect is made obvious on figure 1 and reminds us of the multiple sources of variation. Also, the accounting method presented here is weak because using total length for comparing occurrence numbers or duration doesn't allow the control of speech rate, which can vary depending on the speech modality used and on the language spoken, especially for L2 learners. We thus recommend using the number of uttered creaky syllables upon the total syllable number ratio. We should finally mention that even though creaky voice manifestation in French was not pointed out in the literature, it doesn't mean that this effect is absent, especially because such effects can change over the years, as we underlined it for American English. A comparison to French monolinguals speakers, which is currently in progress, seems thus essential to support our results.

7. Conclusion

This experiment intended to study the phenomenon of creaky voice among American speakers under two different lights: first the impact of a language spoken (English or French) on the appearance of this phenomenon, and second, the impact of the speech task (reading task or spontaneous speech) on this phonatory register. The specificity of this study lies in the fact that the speakers, who are learning French as a second language, are confronted to an unconscious phonatory choice. The results confirmed the hypothesis according to which a differentiated use is manifest and more creaky voice is expected in English and in spontaneous speech for our American speakers learning French. This leads to consider that second language learners do pay attention to phonatory settings, to some extent, when they learn a language. This underlines the necessity of integrating phonation settings aspects in the process of L2 learning. According to Esling [17], when a voice quality aspect is present in the native language of the learner, and absent in the second language, it can be an obstacle to his intelligibility and his social acceptance.

8. Acknowledgements

We would like to thank Lise Crevier-Buchman for her proofreading. We would also like to thank our speakers for their kindness and patience during the recordings and Zoe Falk for helping with the translation. Special thanks go to Laurianne Georgeton and Nicolas Audibert for their precious advice. This work has been supported by the Labex EFL (ANR/CGI)

9. References

- [1] Hollien, H. and Michel, J.F., "Vocal fry as a phonatory register", *Journal of Speech Hearing Research*, 11: 600-604, 1968.
- [2] Hollien, H., Girard, G.T. and Coleman, R.F., "Vocal fold vibratory patterns of pulse register phonation", *Folia Phoniatr (Basel)*, 9: 200-205, 1977.
- [3] Moore, P. and Von Leden, H., "Dynamic variations in the vibratory pattern in the normal larynx", *Folia Phoniatr (Basel)*, 10, 205-238, 1958.
- [4] Timcke, R., Von Leden, H. and Moore P., "Laryngeal vibrations: measurements of the glottic wave", *AMA Arch Otolaryngol.*, 68: 1-19, 1958.
- [5] Kreiman, J., "Perception of sentence and paragraph boundaries in natural conversation", *Journal of Phonetics*, 10: 163-17, 1982.
- [6] Ylitalo, Y. and Hammarberg, B., "Voice characteristics, effects of voice therapy, and long-term follow-up of contact granuloma patients", *Journal of Voice*, 14: 557-566, 1997.
- [7] Ladefoged, P. and Maddieson I., "The sounds of the world's languages", Cambridge, MA: Blackwells, 1996.
- [8] Henton, C. and Bladon A., "Creak as a Sociophonetic Marker", in *Language, Speech and Mind: Studies in honour of Victoria A. Fromkin*, [Ed] Larry Hyman and Charles N. Li, 3-39 London Routledge, 1988
- [9] Pierrehumbert, J., "Prosodic effects on glottal allophones", In *Vocal fold physiology: voice quality control* (Fujimara et Hirano [Ed]), 39-60. San Diego: Singular Publishing Group. 1995.
- [10] Slička, J., "Some physiological correlates to regular and irregular phonation at the end of an utterance", *Journal of Voice*. 20: 171-186, 2006.
- [11] Ogden, R., "Turn-holding, turn-yielding, and laryngeal activity in Finnish talk-in-interaction", *Journal of International Phonetical Association*, 31: 139-152, 2001.
- [12] Ishi, C.T., Ishiguro, H. and Hagita, N., "Proposal of acoustic measures for automatic detection of vocal fry", *Proceedings of Interspeech*, Lisbon, Portugal, 481-484, 2005.
- [13] Carlson, R., Gustafson K. and Strangert E., "Cues for hesitation in speech synthesis", *Proceedings of Interspeech*, Pittsburgh, USA 1300-1303, 2006.
- [14] Podesva, R.J., "Phonation type as a stylistic variable: The use of falsetto in constructing a persona". *Journal of Sociolinguistics*, 11 (4): 478-504, 2007.
- [15] Yuasa, I.P., "Voice quality for young urban oriented upwardly mobile American women?" *American Speech*, 85: 315-337, 2010
- [16] Wolk, L. and Abdelli-Beruh. N.B., Slavin D., "Habitual use of vocal fry in young adult female speakers", *Journal of Voice*, 26 (3): 111-116, 2012.
- [17] Esling, J., "The identification of Features of Voice Quality in social groups", *JIPA* 8: 18-23, 1978.
- [18] Lehiste, I., "Sentence boundaries and paragraph boundaries-perceptual evidence", In: *The Elements: A Para-Session on Linguistic Units and Levels*. Chicago Linguistics Society; 99-109, 1979.
- [19] Sapir, E., "Speech as a personality trait", *American Journal of Sociology*, 32: 892-905, 1927.
- [20] Knowles, G., "The nature of phonological variables in Scouse » In P. Trudgill [Ed], *Sociolinguistic Pattern in British English*, London, Edward Arnold, 80-90, 1978.
- [21] Stuart-Smith, J., "Glasgow: accent and voice quality" In P. Foulkes and G. Docherty [Eds], *Urban voices: accent studies in the british Isles*, London: Arnold, 201-22, 1999.
- [22] Dilley, L., Shattuck Hufnagel S. and Ostendorf M., "Glottalization of word-initial vowels as function of prosodic structure", *Journal of Phonetics*, 24: 423-444, 1996.
- [23] Ingle, J., "Pacific Northwest Vowels: A Seattle Neighborhood Dialect Study", Paper presented at the 149th Meeting for Acoustical Society of America, Vancouver, B.C., May 16-20, 2007.
- [24] Gottliebson, R.O., Lee L., Weinrich B., and Sanders J., "Voice problems of future speech-language pathologists", *Journal of Voice*, 21: 699-704, 2007.
- [25] Böhme, T., Shattuck-Hufnagel, S. "Do listeners store in memory a speaker's habitual utterance-final phonation type?" *Phonetica*, 66: 150-168, 2009.
- [26] Honikman, B., "Articulatory Settings", In D. Abercrombie, D.B. Fry, P.A.D. MacCarthy, N.C. Scott and J.L.M. Trim [Ed], In honour of Daniel Jones, pp. 73-84. London: Longmans, 1964.
- [27] Kane, J., Drugman, T. and Gobl, C., "Improved automatic detection of creak", *Computer Speech and Language*, 27 (4): 1028-1047, 2012.
- [28] Laver, J., "The Phonetic Description of Voice Quality", Cambridge, UK: Cambridge University Press, 1980.
- [29] Redi, L. and Shattuck-Hufnagel S., "Variations in the realization of glottalization in normal speakers". *Journal of Phonetics*, 29: 407-429, 2001.