

Perceptual level of intonation in whispered voice

Géraldine Vercherand

CLILLAC - ARP, UFR Linguistique, Université Paris Diderot, France

<https://doi.org/10.36505/ExLing-2011/04/0037/000206>

Abstract

In this paper, I propose to analyze the prosodic realization of two modalities (declarative and interrogative) through perception study of whispered voice in French. I present results of two modality identification tasks. Results of identification test based on natural stimuli show a good identification rate of the modality. The values of the duration of final syllable seem to be the main acoustic cue that is relevant in the identification of modality. I present also the results of modality identification task based on re-synthesized stimuli allowing to control the role played by duration parameter. Results of this second experiment show modality identification depends on perceptual level of the duration.

Key words : whispered speech, intonation, perception, acoustic parameters.

Introduction

We know that whisper implies no vibration of the vocal folds, provoking thus the absence of fundamental frequency. However intonation continues to be recognized in whisper.

Most of the works on whisper concern its physiological and acoustical aspects (glottal characteristics (Tartter 1989) laryngeal movements (Bonnot 1987) and consonant closure duration (Higashikawa 1996), where generally the exploited material for the analysis consists of isolated words or non-words (Higashikawa 1996). Very few studies deal with the prosodic realization in whisper (Meyer-Eppler 1961) or the differences between voiced speech and whisper in larger items than words such as sentences (Faraco 1983 for French).

My study tries to check results obtained on a corpus more controlled (than Faraco 1983) and to measure the importance of acoustic parameters. So for this study, I will start with 2 hypotheses:

- a) each parameter (values of formants, duration, speech rate and intensity) contributes to the detection of modality ;
- b) rhythmic unit is important to the realization of acoustic parameters.

Natural stimuli

The corpus

The French corpus consists in 11 phrases pronounced as declarative and interrogative. The length was incremented from 2 syllables to 12 syllables, 1 syllable incremented each time.

Table 1 : Corpus (Combination of an element per column).

Elle (1,2)	mange	Ø (1)
Anna (3,4,5,6)		bien (2,3)
Annabelle (7,8,9)		une noix (4)
Annabella (10,11)		un melon (5)
		des ananas (6,7)
		une noix du Ghana (8)
		un melon du Ghana (9,10)
		des ananas du Ghana (11)

Perception test

This perception test based on natural stimuli aims a) to determine if modality is recognized in whispered speech, b) to see if the sentence length (number of syllables) has an effect on the rate of recognition of modality and c) to understand how modality is realized and to determine which acoustics phenomena are decisive.

Perception test results

The modality is fairly recognized in whispered speech. We had 78% response as result. The ANOVA test shows statistically significant effect by length on identification rate ($F(10,198)=5,421, p<0.0001$).

Confirming my hypothesis, the perception rate correlates with the phrasal length (Interrogative : 2 to 5 syllables 67,5%, 6 to 8 syllables 79,5%, 9 to 12 syllables 89,2% ; Declarative : 2 to 5 syllables 68,5%, 6 to 8 syllables 79%, 9 to 12 syllables 82,75%).

Analyze

A total of 6 items, 3 interrogative and 3 declarative were less recognized by all listeners.

- *declarative sentences*. The value of speech rate and final duration are different from the average expected. Duration seems to be the most important parameter.

- *interrogative sentences*. Several parameters seem to be involved, as too slow speech rate, duration of the last syllables too long and a dynamic intensity too high.

Resynthesized stimuli

Perception study based on natural stimuli shows that the value of a parameter seems to compensate the other and vice versa it could alter interpretation. So I choose to vary the duration of the last syllable in order to verify its influence on the identification scores. I hypothesize that increasing the duration of the last syllable compensates the absence of other acoustic parameters.

The corpus

I select three declarative sentences of test corpus based on natural stimuli. These sentences are representative of the number of syllables (3 syllables, 7 syllables, 12 syllables).

I built a continuum of time on the last syllable. The continuum is composed by 5 steps. The synthesized stimuli were generated from the third stimulus of the continuum (Stimulus 1 -80ms, S2 -40ms, S4 +40ms, S5 +80ms). The corpus is composed of 5 stimuli per each sentence. Each stimulus appears three times in the test, representing a total of 45 stimuli.

Perception test

The task of the test is to identify the modality. The listeners are asked to choose among three possibilities: “declarative”, “interrogative” and “I do not know”. 12 French listeners participate. The listeners take first a pre-test to familiarize themselves with the type of stimuli. The stimuli are presented randomly.

Perception test results

Table 3 : Identification scores of modality with synthesized stimuli - in row stimuli, in column responses (%).

Sentences	Simuli	Declarative	Interrogative	I do not know
Short	1	61,11	36,11	2,78
	2	41,67	58,33	0
	3	16,67	83,33	0
	4	8,33	91,67	0
	5	19,44	77,78	2,78
Medium	1	88,89	11,11	0
	2	94,44	2,78	2,78
	3	86,11	8,33	5,56
	4	80,56	13,89	5,56
	5	69,44	27,78	2,78
Long	1	41,67	50	8,33
	2	44,44	55,56	0
	3	25	69,44	5,56
	4	27,78	61,11	11,11
	5	25	66,67	8,33

More duration of the last syllable increases, more listeners identify the sentence as interrogative. An ANOVA test shows a significant effect of treatment duration ($F(4,15)= 1,538E15$, $p<0,0001$), a significant effect of sentence type ($F(2,15)=1,551E16$, $p<0,0001$) and an interaction between the two independent factors ($F(8,15)=3,143E15$, $p<0,0001$).

Thus increased duration of the last syllable influence significantly the recognition rate. The level of duration to identify modality seems to depend on the length of the sentence. The length of the sentence influences the recognition rate and interacts with changes in duration.

Discussion and conclusion

Through this study on whispered speech, it shows that each parameter (duration, speech rate, intensity and formant) involved in the perception of the two modalities patterns, the declarative and interrogative. The length of the sentences parameter is important. It seems that the rhythmic factor plays an important role, since a number of syllables is necessary to allow the parameters of intensity and duration to differentiate sufficiently and to allow perception of different modalities. In addition, this study shows that the prominence of an acoustic parameter compensates for non-prominence of others in the detection of prosodic phenomena studied. However, the required level seems to depend on the length of the sentence. Further studies are underway to explore this question and to widen of perspectives to other modalities.

References

- Bonnot J-F.P. 1987. Electromyographical Correlates of Shouted and Whispered Voice. Actes du 11ème Congrès International des Sciences Phonétiques, volume 4, 20-23, Tallinn, Estonie, 1987.
- Faraco M. 1983. Comparaison des intonations affirmatives et interrogatives en voix normale et voix chuchotée, Thèse de 3ème cycle, Université d'Aix en Provence.
- Higashikawa, & al. 1996. Perceived Pitch of whispered voice—Relationship with formant frequencies: a preliminary study. *Journal of Voice* 10.2: 155-158.
- Meyer-Eppler, W. 1957. Realization of prosodic features in whispered speech. *JASA* 29, 104-106.
- Tartter V.C.1989. What's in whisper? *JASA* 86 (5), 1678-1683.