

“Deep and raspy” or “High and squeaky”: a cross-linguistic study of voice perception and voice labeling

Begoña Payá Herrero

Ludwig Maximilians Universität, Germany

<https://doi.org/10.36505/ExLing-2008/02/0028/000087>

Abstract

Vocal correlates for identity markers were sought by interviewing 60 candidates of three nationalities and by analysing their average pitch, intensity, speech rate and speech length. National differences were found in pitch and specially speech rate but no clear vocal correlates for the personality parameter of introversion/extroversion were found. An analysis of the voice labels given by the candidates showed that there are common voice stereotypes for sexy voices and unpleasant voices, despite some minor national differences in the use of labels such as *sweet* or *nasal*. Finally, a perception test proved that there is a high-interrater consistency in judgements about voice and personality and that basic emotions and attitudes are well recognised, irrespectively of the language knowledge.

Key words: vocal correlates, identity markers, voice perception, voice stereotypes, voice labels.

Introduction

It is through the combination of *voice dynamics* aspects and *voice quality* aspects (Abercrombie 1967:7) that indexical information about the speaker's identity is conveyed through voice. Indexical information can be said to be either *biological*, *psychological* or *social* (Laver, 1991: 154) or from the perspective of the vocal cues that give this information, we can speak about *physical markers*, *psychological markers* and *social markers* (Laver & Trudgill, 1979:2). When it comes to judgments people make about the three kinds of indexical information mentioned when hearing a voice, *psychological* and *social information* judgments tend to be less accurate, just because these are the ones that are culturally based, arbitrary and learnt. Yet these are the ones upon which I will be focusing my research, since they are the “fingerprints” of our emotions, attitudes and personality. A starting point for my project was to collect voice labels that non-experts use that would include young female speakers of Spanish, German and English.

Interviews

My target group was made up of female speakers between 20 and 30 years of age: 20 Germans, 20 Spanish, and 20 Americans. Face-to-face individual interviews were carried out in each of their mother tongues and they had to answer a total of 11 questions.

Voice labels

Using Laver's distinction (1974), voice labels provided by candidates in their answers were mainly classified into **phonetic labels** and **impressionistic labels**. The following interesting observations were made:

- In describing their own voice and sexy voices, there was a clear tendency for candidates of the three nationalities to choose *phonetic labels* more often than *impressionistic labels*, with *voice dynamics* and pitch labels, being the most frequent ones.
- In the description of unpleasant voices there was again a clear preference amongst the three nationalities for *phonetic labels*, but this time *voice dynamics* labels and *voice quality* labels were equally preferred. Two adjectives were found to be the most frequent in all three nationalities: *high* and *squeaky*, which were the stereotypes for unpleasant voices.
- A stereotypical sexy voice seems to be the *deep and raspy* kind in the case of all the languages studied. This then would be, at least for western culture, to which the languages studied belong, the perceived factors pointing to sexiness in voice. This same idea can be found in Rodero's (2001: 8-9) study of pleasant voices in radio news.
- Another interesting observation is the perception of *nasality* as unpleasant only for the German and American candidates. *Nasal* was only used once by a Spanish speaker, whereas, it was used by three different Germans and by five different Americans. This seems to show that only in the phonological systems where nasality is strongly present (at least here for American English and German), can this aspect be perceived as unpleasant.

Voice, nationality and personality

Candidates interviewed also answered a personality test, the so-called Jung Myers Briggs Typology test in order to classify them according to the introversion/extroversion parameter of personality. All the voice samples from the interviews were then analyzed phonetically with Praat (length, intensity, frequency) and the average speech rate in SPM was calculated. The phonetic results were then contrasted in terms of nationalities and personality groups (introverted vs. extroverted) to see if there were on the one hand big national differences and on the other hand to search for clear correlations between specific vocal cues and the degree of introversion/extroversion. Certain national differences were found in speech rate (spanish women being the fastest speakers) and in pitch (spanish women had the lowest average pitch, german women the highest), but not in intensity. See figures 1 and 2.

However, no obvious correlation could be found between the degree of introversion/extroversion and particular trends in the use of voice cues (such as expected high speech rate for extroverts or lower intensity for introverts). Extroverts did speak longer in average throughout all nationalities but not all personality groups displayed general tendencies in pitch, intensity or speech rate. This seems to indicate that introversion or extroversion is not clearly manifested by a single voice parameter.

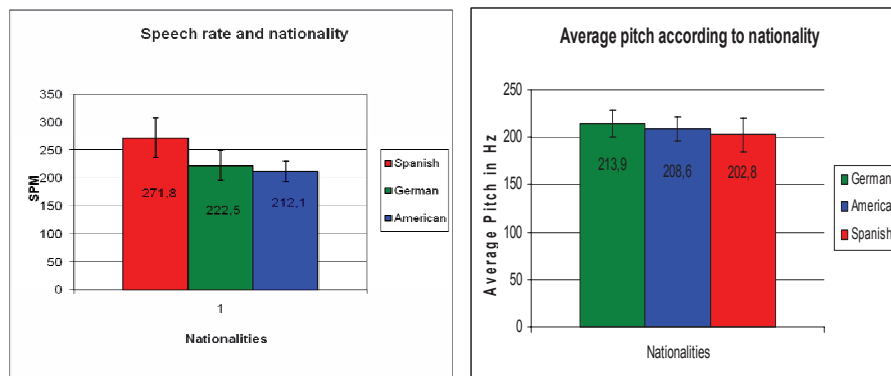


Figure 1. Speech rate and nationality. Figure 2. Pitch and nationality

Perception test

The next step was to move on from the phonetic analysis and the label classification to the perception of voice by non experts. The perception test was devised by using differentiated female voices from the interviews and acted emotions and attitudes as stimuli and the labels from the classification in a bipolar five point scale, inspired by Osgood's (1964) semantic differential technique. A total of four different groups of voice judges took the test: some understanding the target language other not. All the answers given by a total of 40 voice judges were then classified in an Excel table that gave the average results for each stimulus heard.

In general, in terms of judgments on voice dynamics and voice quality, the results showed that tone and pitch are the easiest recognized. Regarding the judgments on personality, there was a correlation between voices felt as self-confident/strong/adult and extrovert. Voices perceived as emotional were also felt to be aggressive and squeaky. Similarly there was a correlation between fast and active voices. In terms of the judgments on synesthetic labels, voices perceived as high, were all judged to be bright and light, and voices seen as warm were also in general soft. This all proves that there is high interrater consistency in the answers.

Moreover, the initial stereotypes found in the interviews claiming that in general high voices are unpleasant, whereas deeper voices tend to be attractive, was also confirmed in the perception test.

When it came to the perception of emotions and attitudes, most candidates, irrespectively of their knowledge of the language heard, made right guesses.

This then supports the idea that attitudes and emotions are transmitted more by different vocal cues than by the content level of language, at least for the languages studied.

Conclusions

Within a relatively small scope of cross-linguistic samples, the present study has perhaps firstly shed some light upon how voice perception is mirrored by voice labels. There are voice stereotypes that are not purely language dependant, but rather, common to the Western culture they belong to. However, a small number of language-dependant voice labels have also been found in this study. Secondly, the perception test carried out echoes many other projects that have stated how voice is a key transmitter of indexical information, by showing that there is something we could call “indexical stereotypy of voice”²: in other words, that people often do make interpretations of indexical information based on vocal cues, whether what is inferred is accurate or not. Thirdly, the crosslinguistic perception test has helped support the idea that it is the phonetic level of language -and not the content level- that conveys most of the indexical information, and that it helps people make accurate guesses about emotions and attitudes.

Acknowledgements

I thank my two PhD tutors, Dr. Pennock and Prof. Schulze for their supervision and I thank the Mutua Madrileña for their economic support.

References

- Abercrombie, D. 1967. *Elements of General Phonetics*. Chicago: Aldine.
- Laver, J. 1974. Labels for voices. *Journal of the International Phonetic Association* 4: 62-75
- Laver, J. and Trudgill, P. 1979. Phonetic and linguistic markers in speech. In Scherer, K.R. and H.Giles, H. (eds.), *Social markers in speech*, 1-32. Cambridge: Cambridge University Press.
- Laver, J. 1991. *The Gift of Speech. Papers in the Analysis of Speech and Voice*. Edinburgh: Edinburgh University Press.
- Rodero, E. 2001. El tono de la voz masculina y femenina en los informativos radiofónicos: un análisis comparativo. Valladolid: Congreso Internacional de Mujeres, Hombres y Medios de Comunicación. Retrieved 11/03/2007 from www.bocc.ubi.pt/pag/rodero-emma-ono-voz-femenina.pdf
- Osgood, C. 1964. Semantic Differential Technique in the Comparative Study of Cultures. *American Anthropologist* 66/3, 171-200.