

Vocal stereotypes as characterisation in animated films

Alice Crochiuia^{1,2}, Anders Eriksson², Sandra Madureira¹

¹PUC-SP, Brazil

²Stockholm University, Sweden

<https://doi.org/10.36505/ExLing-2024/15/0008/000633>

Abstract

Humans have long associated voice with personality and physical traits, with consistent vocal stereotypes evident across cultures. This paper examines how vocal stereotypes influence character perception in animated films, focusing on the multilingual dubbing of *Zootopia* in English, Brazilian Portuguese and Swedish. Animated media provides a unique lens for studying these correlations, as character voices are intentionally matched to stereotypical features, and dubbing allows cross-language comparisons. This study considers Ohala's Frequency Code, which links vocal pitch to size and dominance, rooted in both human and animal behaviour. The analysis includes acoustic parameters, voice quality, and listener perceptions. Results show the use of higher f_0 to portray smaller characters, in contrast with larger ones, which display lower f_0 across all languages.

Keywords: dubbing, animated films, vocal stereotypes, voice quality, voice analysis.

Introduction

Vocal stereotypes are present not only in daily, face-to-face interactions but also permeate media and entertainment, as they function as effective shortcuts to establish a character's features to the audience. Animated films and TV series provide a unique lens for investigating vocal stereotypes, as they require voices to be matched to the physical and psychological features stereotypically correlated to them. In addition, animated features are often dubbed in different languages for international distribution, thus allowing for the comparison of how the same character sounds in various languages.

This paper outlines materials and methods chosen for a multilingual study of how vocal features are used to reflect physiological, psychological and social cues in an animated film. It leverages theories like the Frequency Code (Ohala 1984), which describes the association of f_0 and body size and extends those correlations of size to social behaviours such as dominance and submission. The paper also presents the results and discussion of the experiments and analyses carried out so far in the study.

Material and methods

Material

The material for this study comprises the dialogue lines of the four main characters of the animated film *Zootopia*, in its original version in English and its dubbings in Brazilian Portuguese and Swedish. These characters are Judy Hopps (a female bunny), Nick Wilde (a male red fox), Chief Bogo (a male African buffalo) and Assistant Mayor Bellwether (a female sheep).

For the analyses, one audio sample was produced for each character in each language, edited together from dialogue lines from different scenes in the film. For Bellwether, two audio samples were produced, as the character is revealed as the film's villain and presents two different personas during the narrative. Each audio sample was around 27 seconds in length.

Perceptual experiment

The perceptual experiment for the audio samples in Brazilian Portuguese was conducted entirely online, with 77 Brazilian Portuguese native speakers, male and female, aged between 20 and 50. The judges were asked to listen to the five audios to evaluate the speaker using 14 bipolar scales speakers' size, age, temper, attitudes, character, and social and vocal features. This was done using a sliding tool to indicate the scale point they felt best fit the character, which was translated into a score from 0 to 100.

The descriptors chosen for the test were based on features displayed by the characters in the film and on descriptions given by the writers, directors, and voice actors in interviews, discussions, and promotional material.

Acoustic analysis

A modified version of the script Prosody Descriptor Extractor (Barbosa 2020) for Praat (Boersma & Weenink 2021) was used for the acoustic analysis of the Brazilian Portuguese, Swedish and English samples. The script automatically extracts measures of parameters related to f_0 , intensity, and long-term spectrum, among others.

Perceptual voice quality analysis

Voice quality analysis of the Brazilian Portuguese samples was carried out by seven phoneticians, using the VPA protocol (Laver & Mackenzie Beck 2007).

Results

Table 1 shows selected results from the perceptual experiment with audio samples from the Brazilian Portuguese dubbing. It displays the average ratings of each character on the scales Small – Big, Docile – Aggressive, Submissive – Dominant. Higher values indicate ratings closer to the descriptors listed in the table, and lower values ratings closer to the opposite descriptor. Highlighted in

bold are the strong scores (<25, >75) each character received in these scales. Bellwether 1 represents the character's friendly façade, while Bellwether 2 represents her true scheming personality.

As Table 2 shows, Chief Bogo had the lowest f_0 values across all languages, while Bellwether 1 had the highest.

Bellwether 1 and Bellwether 2 present different voice quality settings in Brazilian Portuguese, as seen in Table 3. Chief Bogo, in addition to low f_0 , also uses voice quality settings that help lower the overall tone of voice even further.

Table 1. Average ratings for the characters in the perceptual experiment.

Characters	Big	Aggressive	Dominant
Bellwether 1	29.97	12.64	44.36
Bellwether 2	24.09	68.99	88.13
Chief Bogo	86.71	86.42	90.69
Judy Hopps	34.40	48.51	68.70
Nick Wilde	53.48	42.81	65.65

Table 2. Mean f_0 values in the three languages.

Characters	English	Brazilian Portuguese	Swedish
Bellwether 1	237 Hz	222 Hz	242 Hz
Bellwether 2	199 Hz	178 Hz	202 Hz
Chief Bogo	109 Hz	127 Hz	124 Hz
Judy Hopps	194 Hz	203 Hz	189 Hz
Nick Wilde	148 Hz	135 Hz	150 Hz

Table 3. Voice quality and vocal dynamics settings of Brazilian Portuguese voices in Zootopia

Characters	Voice quality settings	Vocal dynamics settings
Bellwether 1	Lip Spreading, Whispery Voice	High Mean Pitch, High Mean Loudness
Bellwether 2	Harshness, Extensive Mandibular Range	High Mean Pitch, High Mean Loudness
Chief Bogo	Lip Rounding, Lowered Larynx	Low Mean Pitch, High Mean Loudness
Judy Hopps	Extensive Labial Range, Raised Larynx	High Mean Pitch, Extensive Pitch Range & High Pitch Variability
Nick Wilde	Pharyngeal Expansion, Lax Vocal Tract	Extensive Pitch Range & High Pitch Variability

Conclusions

The use of different voice quality settings for the friendly Bellwether and her conniving persona supports the use of distinct vocal features to signify different psychological features in characters. Bellwether 1's use of the Lip Spreading setting, which often correlates to perception of a smile, also raises the overall tone of a speaker's voice, as it shortens the length of the vocal tract.

The listeners' perception of body size matched the descriptions in Ohala's Frequency Code hypothesis. Chief Boggo, who was accurately perceived as the largest of the characters, had the lowest f_0 across the three languages. In contrast, the two smallest characters, Judy and both iterations of Bellwether, had the highest values (the film inaccurately depicts Bellwether, who's a sheep, as smaller than Nick, who's a red fox).

On the other hand, the perception of aggressiveness and dominance were also influenced by voice quality settings and loudness. The combination of a low f_0 , Lowered Larynx, Lip Rounding and High Loudness in Chief Boggo's voice signalled both aggressiveness and dominance. Bellwether 2, who was perceived as dominant but not aggressive, had a combination of a high f_0 , Harshness, Extensive Mandibular Range and High Loudness.

Acknowledgements

The first author acknowledges grants from CAPES #88887.630904/2021-00 and #88881.689915/2022-01. The second author acknowledges a grant from The Swedish Foundation for International Cooperation in Research and Higher Education (STINT) (IB2015-6488). The third author acknowledges a grant from CNPq #302194/2019-3, and a grant from PIPEq-PUCSP #211463.

References

- Crochiquia, A., Eriksson, A., Barbosa, P. A., Madureira, S. 2022. A perceptual and acoustic study of dubbed voices in an animated film. In Frota, S., Cruz, M., Vigário, M. (eds.), Proc. 11th Intern. Conf. on Speech Prosody, 565-569, Lisbon, Portugal.
<https://doi.org/10.21437/SpeechProsody.2022-115>
- Crochiquia, A., Eriksson, A., Madureira, S., Barbosa, P. A. 2023. Animated film character profile: the roles of voice and lexical content. In Skarnitzl, R., Volín, J. (eds.), Proc. 20th ICPHS 466-47, Prague, Czech Republic.
<https://guarant.cz/icphs2023/862.pdf>
- Liu, W., Zhang, X., Liang, C. 2023. An acoustic study on character voices of dominators and subordinates: A case study on male characters in Empresses in the Palace. In *Frontiers in Communication* 7.
<https://doi.org/10.3389/fcomm.2022.1088170>
- Ohala, J. J. 1984. An Ethological Perspective on Common Cross-Language Utilization of F0 of Voice. In *Phonetica*, 41(1), 1–16.
<https://doi.org/10.1159/000261706>