

Madureira, S.; Andreassa, J.; Brisola, E.; Crochiquia, A.; Fontes, M.; Pires, M.;

Scarpelly, R. (2023)

Pontifical Catholic University of São Paulo

Corresponding author e-mail address: sandra.madureira.liaac@gmail.com

Vocal and Visual Symbolism

Vocal and facial gestures are associated to express meanings in speech productions. These two non-verbal aspects are key factors in spoken communication to reinforce the semantic content, contradict the semantic content, or adding extra information. Their integration can be interpreted as originating from the same embodied mechanisms (Vainio and Vainio, 2021).

This paper maps and comments on the theoretical and methodological bases of the research works which are developed by our research group on vocal, facial, and meaning relations underlying speech expressiveness. These research works comprise the investigation of settings of voice quality to build animation characters and uses of vocal prosody in dubbing (Crochiquia et al, 2020; Crochiquia et al 2022), impressionistic judgments of voice qualities based on semantic descriptor assignments (Scarpelly, Passetti and Madureira, 2022), characteristics of charismatic speech, vocal characteristics of emotional speech (Fontes and Madureira, 2015) and vocal and visual prosodic characteristics of speech, reciting and singing styles (Lomba, Fontes and Madureira, 2017; Madureira, 2018).

In our research group, experimental investigations on how meanings are expressed by speakers and perceived by listeners are carried out with a diversity of methodological tools and methods and a unified gesture-centered theoretical analysis within a metaphorical framework based on the frequency, production, effort and sirenic sound symbolism codes, perceived duration, and crossmodal and synesthetic interactions that language entails (Nobile, 2019).

From the methodological point of view, perceptual analysis of segmental and vocal prosodic elements, perceptual analysis based on semantic descriptor evaluation, acoustic analysis, automatic analysis of facial gestures is applied, and their results correlated by means of Exploratory Multivariate statistical analysis (Husson Pagés and Lê, 2009). Multivariate statistical analysis comprises several methods: PCA, MCA and MFA, FAMD. The latter two are used to correlate quantitative and qualitative variables. The Voice Profile Analysis system (Laver and Mackenzie Beck, 2007) is applied to perform perceptual analysis of voice quality and prosodic elements. Speech segments are analyzed according to their articulatory and acoustic criteria and their characteristics are interpreted in relation to symbolic, iconic and metaphorical sound meaning relations Hinton, Nichols, and Ohala (1994). Another kind of perceptual analysis involve the use of semantic descriptors in perceptual tests applied to get knowledge on the way listeners attribute meanings based on vocal and visual features.

For the acoustic analysis the Prosody Descriptor Extractor script for Praat (Barbosa, 2021) is used. The script computes 30 prosodic parameters related to melodic, rhythmic and voice quality features, generating statistical descriptors and acoustic measures of F0, intensity, long-term spectrum, harmonic to noise ratio, duration of vowels and silence.

Automatic analysis of facial expressions uses the FACs system (Ekman, Friesen, and Hager, 2002) and associates the facial movements described by its analytic unit, the

Action Unity, to emotion expressions. In computer-based facial expression analysis, emotions are detected by computer algorithms that record facial expressions via webcam. Action Unities describe the movements of both the upper and the lower part of the face. Head movements are also described in the FACS system.

For the sake of demonstrating the kind of framework analyzed, a small experiment consisting of producing the same utterance with labial, mandibular and tongue tip/blade settings of voice quality is presented. These settings were chosen because of their visual and vocal features. These settings are seen and heard, and their productions involve contrastive spatial and directional movements (spreading and rounding; opening and closing; fronting and backing). The effects of these contrastive voice quality settings and facial movements on meaning expression are considered from a crossmodal sound-symbolic perspective.

References

- Barbosa, P. (2021) Prosody Descriptor Extractor” [Praat script]. Available: <https://github.com/pabarbosa/prosodyscripts/tree/master/ProsodyDescriptorExtractor>.
- Crochiquia, A., Eriksson, A., Fontes, M. A., Madureira, S. (2020). Um estudo fonético das vozes de personagens do filme Zootopia na dublagem em português brasileiro: o papel dos estereótipos. DELTA. 36, 311. doi: 10.1590/1678-460x2020360311.
- Crochiquia, A., Eriksson, A., Madureira, S., Barbosa, P. (2022). A perceptual and acoustic study of dubbed voices in an animated film. In: Proceedings of Speech Prosody 2022. Lisboa: ISCA, 2022. v. 1. p. 565-569.
- Ekman, P., Friesen, W. V., and Hager, J. C. (2002). Facial Action Coding System (2nd ed.). Salt Lake City, UT: Research Nexus eBook.
- Fontes, M., Madureira, S. (2015) Gestural prosody and the expression of emotions: a perceptual and acoustic experiment. In: Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow: University of Glasgow, v. 1. p. 1-5.
- Gussenhoven, C. (2002). “Intonation and interpretation: Phonetics and phonology,” in Proceedings of the 1st International Conference on Speech Prosody (Aix-en-Provence), 47–57.
- Gussenhoven, C. (2004). The Phonology of Tone and Intonation. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511616983.
- Gussenhoven, C. (2016). Foundations of intonation meaning anatomical and physiological factors. Topics Cogn. Sci. 8, 425–434. doi: 10.1111/tops.12197.
- Hinton, L., Nichols, J., and Ohala, J. (1994). Sound Symbolism. Cambridge: Cambridge University Press.
- Husson, F., Lê, S., and Pagès, J. (2009). Exploratory Multivariate Analysis by Example Using R. London: Chapman and Hall.
- Laver, J., Mackenzie-Beck, J. (2007). Vocal Profile Analysis Scheme -VPAS [handout]. Edinburgh: Queen Margaret University College, Research Centre.
- Lomba, J. A.; Fontes, M. A. S.; Madureira, S. (2017) As Palavras Cantada, Falada e Declamada: Um Estudo Sobre a Relação Entre Canto, Fala e Declamação e seus Efeitos Impressivos. In: Sandra Madureira. (Org.). Sonoridades: a expressividade da fala, no canto e na declamação. 1ed. São Paulo.
- Madureira, S. (2018) Brazilian Portuguese rhotics in poem reciting. In: Mark Gibson; Juana Gil. (Org.). Romance Phonetics and Phonology. 1ed. Oxford: Oxford University Press, v. 1, p. 191-215.
- Nobile, L. (2019). Introduction: Sound symbolism in the age of digital orality. A perspective on language beyond nature and culture. Signifiances (Signifying), 3, XXXVI–LXVIII. doi: 10.18145/signifiances.v3i1.248.

Ohala, J. J. (1994). "The frequency codes underlies the sound symbolic use of voice pitch," in Sound symbolism, eds. L. Hinton, J. Nichols, and J. J. Ohala (Cambridge: Cambridge University Press) 325–347. doi: 10.1017/CBO9780511751806.022.

Scarpelly, R., Passetti, R., Madureira, S. Avaliação impressionística e fonético-descritiva de qualidades de voz: convergências, divergências e contextos de aplicação forense. In: Proceedings of the II Brazilian Conference on Prosody, 2022, Belo Horizonte online. 1. p. 120-133.

Vainio L, Vainio M. Sound-Action Symbolism. *Front Psychol.* 2021 Sep 14;12:718700. doi: 10.3389/fpsyg.2021.718700. PMID: 34594278; PMCID: PMC8476841.