



OPEN ACCESS

EDITED AND REVIEWED BY
Xiaolin Zhou,
Peking University, China

*CORRESPONDENCE
Plinio Almeida Barbosa
✉ pbarbosa@unicamp.br

RECEIVED 13 July 2023
ACCEPTED 21 July 2023
PUBLISHED 04 August 2023

CITATION
Barbosa PA (2023) Editorial: Science,
technology and art in the spoken expression of
meaning. *Front. Commun.* 8:1257945.
doi: 10.3389/fcomm.2023.1257945

COPYRIGHT
© 2023 Barbosa. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Editorial: Science, technology and art in the spoken expression of meaning

Plinio Almeida Barbosa*

Department of Linguistics, State University of Campinas, Campinas, Brazil

KEYWORDS

speech research, vocal aesthetics, prosodic, prosodic meanings, speech technology

Editorial on the Research Topic

[Science, technology and art in the spoken expression of meaning](#)

The set of papers in this Research Topic on Science, Technology and Art in the Spoken Expression of Meaning covers a broad spectrum of pragmatic meanings conveyed by voice quality, speech prosody and body gestures in the contexts of human-human and human-machine interactions.

In human-human interaction, the acoustic characteristics of attitudes are explored together with the prosodic attributes of oral poetry and storytelling in different African languages, and the shapes of pragmatically distinct meanings of questions in French and Italian. Multimodality is investigated for the role of facial features in clear speech and emotions.

Human-machine interaction is explored in such manner as to pave the way for a more anthropomorphic vocal expression on the part of machines including the complex perception of racial components, charisma and degrees of arousal in voice.

These pieces of work present research in languages as diverse as Mandarin, Wuxi, French, Italian, English, Yorùbá, Anyi, Ega, Estonian, and Brazilian Portuguese.

The paper by [Ji et al.](#) investigates the encoding of speaker (un)confidence in Wuxi dialect vowels. They show that this propositional attitude is revealed by segmental parameters such as F1 and F2, and by statistical descriptors of prosodic parameters such as F0, intensity and duration. The expression of these parameters interacts with specific tones, namely flat vs. counter tones. Their findings shed new light on the mechanisms of segmental and prosodic encoding of speaker confidence at the vowel level.

The paper by [Liu et al.](#) investigates the character voices of leading male characters in the TV series *Empresses in the Palace*. The authors found that the subordinated characters usually adopt a higher pitch or breathy voice whereas the dominant characters use a lower pitch or modal/creaky voice. In addition, cepstral peak prominence (CPP), F0, and H1-A3 are the key acoustic indicators to distinguish character voices. These results have impact for the entertainment industry, such as the choice of voices for characters in animated films.

Two papers explore the use of the voice for oral poetry and storytelling in African languages. The one by [Akinbo et al.](#) is an acoustic study of the vocal expressions of two genres of Yorùbá oral poetry. An original poem in speech mode was acoustically analyzed and the results showed that cepstral peak prominence (CPP), the Hammarberg index, and the energy below 500 Hz in voiced sounds distinguish

the two genres of oral poetry and speech but are not as reliable as F0 height and vibrato. The paper by [Gibbon](#) studied the rhythm of storytelling in two Niger-Congo tone languages, Anyi and Ega. He showed that the interpretations of the rhythm patterns he found were related to turn interaction types, speech registers and social roles, distinct narration styles, and the genre difference between interactive narration and more formal narratives.

The work by [Cresti and Moneglia](#) reveals the correlations between melodic contours and question speech acts in Italian and French. They explore a classification of question illocutionary types present in two parallel corpora of informal speech, one in Italian and the other in French. Yes/no questions were found not only to end in canonical rising contours, but also decreasing ones (26% in French and 36% in Italian), representing 37% of all questions in French and 39% in Italian. A bit <10% of utterances are questions compared to declarative utterances, which are >50% in both languages. Partial questions are as frequent among questions as 26% in French and 38% in Italian. The authors also highlight the importance of questions performed through certain illocutionary types (tag-questions, alternative questions, and double questions) representing 5% in French and 9% in Italian.

The paper by [Garg et al.](#) investigates the facial cues associated with clear vs. conversational speech in Mandarin. By comparing movements of the head, eyebrows and lips associated with these two speech styles in Mandarin tone articulation, they examined the extent to which clear-speech modifications involve signal-based exaggerated facial movements or code-based enhancement of linguistically relevant articulatory movements. They found that head, eyebrow, and lip movements correlate with pitch-related variability. They also found, for all the four Mandarin tones, longer duration and greater movements of the head, eyebrows, and lips in clear speech in contrast with conversational speech.

The paper by [Madureira and Fontes](#) considers Laver's VPA settings under a sound-symbolic and synesthetic perspective by focusing on the auditory impressions these settings have on listeners' attributions of meaning and associations between vocal and visual features related to the expression of six basic emotions. Their results provide evidence of existing links between sound and meaning which has a relation to the biological codes proposed in the literature, phonetic metaphors, and the vocal and facial gestures involved in emotion expression.

The work by [Holliday](#) analyzed the acoustic properties and racialized judgments of four voices of the Siri assistant. A large set of American English listeners responded to questions about the synthetic speaker's sociolinguistic characteristics and personal features. Her evaluation showed that two of the voices were significantly more likely to be rated as belonging to a black speaker than the other two. Additionally, one of the

voices judged as belonging to a black speaker was judged less competent, less professional, and funniest. VQ measures such as mean F0 and H1–A3c significantly affect the listeners' ratings of the voices and are correlated with perceptions of pitch and breathiness, respectively.

Another work on synthetic voices is the one by [Pajupuu et al.](#) which evaluated the likability of calm and energetic audio advertising styles transferred to Estonian synthesized voices. They used a corpus of advertisements created out of the reading of one text in the two advertisement style to show that not only the calm style was preferred, but also that it differed from the energetic one in acoustic features related to a lower, quieter, and more sonorous voice and a more neutral speaking style.

Finally, the paper by [Fucinato et al.](#) showed that charismatic speech features in robot instructions impacts in both team creativity and performance. For doing so, they compared the performance of the teams' activities upon reception of instructions from the robot in a "charismatic" speaking style vs. a neutral way of speaking. The results show that when the robot's speech is based on charismatic characteristics, it is significantly better at enhancing team creativity and performance.

The papers introduced here explore expressive uses of non-verbal language features. These features are multimodal in nature and play a very important communicative role. The main assumption underlying these studies is symbolism, which is manifested in vocal and body language gesture.

Author contributions

PB: Writing—original draft, Writing—review and editing.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.