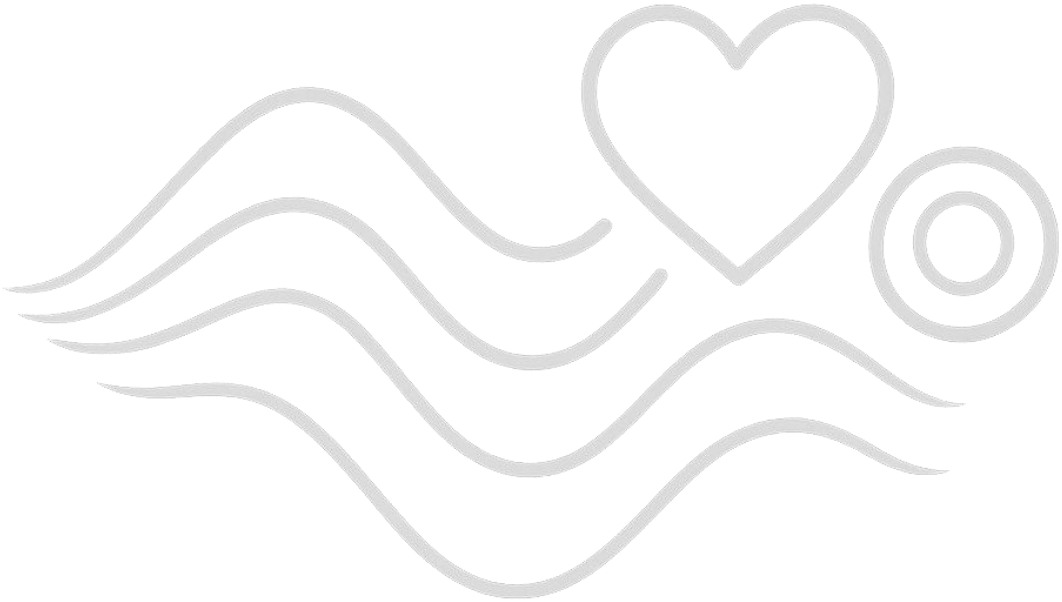
University of Brighton

LEVERHULME TRUST

Swiss National Science Foundation

UNIVERSITÉ DE GENÈVE
CENTRE INTERFACULTAIRE EN SCIENCES AFFECTIVES

BOOK OF ABSTRACTS



University of Brighton
LEVERHULME TRUST
Swiss National Science Foundation
UNIVERSITÉ DE GENÈVE
CENTRE INTERFACULTAIRE EN SCIENCES AFFECTIVES

Beyond mechanics:

Throat vibrations and the embodiment of emotion in voice signals

Leonardo Caravolo

University of Geneva

leonardo.ceravolo@unige.ch

Social communication entails complex dynamics through which one can infer and understand the affective state of one another, for example through the automatic extraction of subtle acoustic features from speech and voice signals. As humans, emotions are therefore embodied and a particularly good blueprint for such mechanism is the occurrence of throat vibrations upon voice production. These vibrations are transmitted through tissues and originate from the oscillations of the vocal folds. Whether these vibratory signals can be interpreted or used consciously by human participants and automatically by the brain remains elusive. In a series of three studies, we investigated whether human participants could accurately evaluate the bodily localization of vibrations emitted through vocal emotion production (Study 1, behavior & throat accelerometry); whether specific brain regions would correlate with such emitted vibrations (Study 2, functional MRI & throat accelerometry); finally, whether induced throat vibrations could influence vocal emotion-based decision-making as well as clarify the brain dynamics involved (Study 3, electroencephalography & throat vibrator). Our results confirmed the hypothesized: accurate localization and intensity assessment of vibrations occurring during vocal emotion production, especially in the upper torso and throat region (Study 1); brain correlates of emitted vibrations upon emotional voice production in the voice-sensitive superior temporal cortex, insular, motor, and somatosensory cortices illustrating bodily self-consciousness and matching voice production networks (Study 2); and biased decisions on the explicit categorization of emotionally-ambiguous voices, especially with induced angry vibrations, relying on a 350-400ms post-onset positivity in a fronto-central cluster of electrodes and originating in the motor and somatosensory cortices as well as in the supplementary motor area (Study 3). Taken together, our data suggest a crucial role of voice-related vibrations both in the perception and production of affective voice signals, and refine our understanding of human social communication through the embodiment of emotion in voice.

How does roughness of human screams affect episodic memory?

Monika Riegel

University of Geneva

monika.riegel@unige.ch

Screams are emotionally salient alarm communications that signal imminent danger and evoke immediate responses. Human screams are characterized by an acoustic attribute of roughness, which reliably increases the subjective aversiveness, captures attention and accelerates reactions, as modulated by the primary auditory cortex and the amygdala in the brain (Arnal et al., 2015, 2019). Despite a well-established effect of salient and aversive signals on different aspects of episodic memory (Riegel et al., 2023; Rimmele et al., 2011), how the roughness of human screams affects physiological arousal and subsequent memory outcomes, remains unknown.

We conducted four experiments combining behavioral (Exp. 1–4) and physiological measures (Exp. 3) using an emotional memory paradigm with acoustically controlled stimuli (roughness, pitch). Participants encoded neutral objects presented in colored frames, each followed by either a scream or a neutral vocalization, and memory for items (objects) and context (frame colour) was subsequently tested. To examine effects on consolidation, memory was tested immediately after encoding (Exp. 1) or after 24 hours (Exp. 2). Moreover, pre- versus post-encoding effects of screams were compared in Exp. 4. Subjective arousal was measured across all experiments, along with electrodermal activity, heart rate, and pupil dilation (Exp. 3).

In line with prior findings, screams were rated as more arousing and negative than neutral vocalizations at encoding and immediate retrieval, but this affective difference faded after 24 hours. Contrary to our hypotheses, item memory was not enhanced by the screams, regardless of a delay before retrieval. However, context memory was consistently impaired if followed, but not preceded, by human screams vs. neutral vocalizations, both tested immediately and after 24 hours. These results indicate that the roughness of human screams selectively and persistently affect certain aspects of episodic memory.

References

Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A. L., & Poeppel, D. (2015). Human Screams Occupy a Privileged Niche in the Communication Soundscape. *Current Biology, 25*(15).

Arnal, L. H., Kleinschmidt, A., Spinelli, L., Giraud, A. L., & Mégevand, P. (2019). The rough sound of salience enhances aversion through neural synchronisation. *Nature Communications*, *10*(1), 1–12.

Riegel, M., Granja, D., Amer, T., Vuilleumier, P., & Rimmele, U. (2023). Opposite effects of emotion and event segmentation on temporal order memory and object-context binding. *Cognition and Emotion*, 1–19.

Rimmele, U., Davachi, L., Petrov, R., Dougal, S., & Phelps, E. A. (2011). Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. *Emotion, 11*(3), 553.

The music of speech:

Relevance, prosody and melody

Joe Reynolds

University of Brighton

j.reynolds10@uni.brighton.ac.uk

"Since intonation is synonymous with speech melody, and melody is a term borrowed from music, it is natural to wonder what connection there may be between music and intonation."

(Bolinger 1986)

The connection between speech prosody and music introduced in the above epigraph is often remarked upon informally and arguably the most compelling parallel lies in the way each modality conveys emotion. Darwin (1871) proposed a shared evolutionary origin, and neuroimaging studies have revealed similar neural networks are recruited in their processing. However, much work is required to address the connection as systematically as it deserves.

Language is a code and words are decoded for interpretation, but a melody does not encode anything at all. Relevance Theory (Sperber & Wilson 1995) is a pragmatic theory in which communicated stimuli are interpreted inferentially. Stimuli may be linguistic but may also be natural or non-verbal. My work explores the extent to which listeners may use the same or related processes in the interpretation of music.

I use the ethological distinction between natural signs – natural communicative phenomena interpreted by inference – and natural signals, interpreted via decoding and inferential processes, to demonstrate resemblances between speech prosody and music. In verbal communication, trembling tends to convey fear. In music, it has become stylised in the vibrato technique to make music 'more emotional' (Seashore 1932). Relevance theory already offers a substantial account of the interpretation of speech prosody. In this talk I present examples that demonstrate clear parallels with music and argue that relevance theory offers insights into the interpretation of both.

References

Bolinger, D.L. (1986). *Intonation and its parts: Melody in spoken English*. Stanford University Press.

Darwin, C. (1871). *The descent of man*. John Murray.

Seashore, C.E. (1932). *The vibrato*. University of Iowa Press.

Sperber, D. and Wilson, D. (1995). *Relevance: Communication and cognition*. Basil Blackwell.

When music "tells time":

Affective meaning and temporal inference

Omayma Rezk

Capital University, Cairo

omayma_abdelfattah@arts.helwan.edu.eg

Research suggests a relation between music, affect, and time perception. The perception of music can serve as a memory-based affective cue that is tied to associative memory of time so listeners report that a certain piece of music "feels like" morning, or "feels like" night, even without any previous temporal reference. From the perspective of memory, temporal concepts are grounded in bodily states, affect, and bodily sensorimotor experiences (Barsalou, 2008; Lakoff & Johnson, 1999). I propose that music perception interacts with previously experienced time experiences by activating similar embodied temporal simulations. My study builds on an empirical foundation that music affects, and even alters, subjective time perception when both arousal and emotional valence interact, and this influences the way time is experienced (whether it passes quickly or slowly) (Droit-Volet et al., 2013). The idea is that musical cues influence the listener's subjective experience of time in terms of temporal categories of morning, night or even seasons, which suggests that time perception is affect-dependent. If certain musical patterns/cues make temporal assumptions more salient, their interpretation aligns with the inverse relation between processing effort and cognitive effects proposed by Relevance Theory (Sperber & Wilson, 1995). Music, then, acts as a stimulus that guides interpretation rather than propositionally encodes it (Wharton, 2009), by activating embodied and memory-based temporal representations. The activation of embodied memory increases the accessibility of temporal assumptions, which are cognitively selected to achieve optimal relevance in the interpretation of music. To infer temporal meaning in music, I propose a relevance-theoretic account of affective meaning in relation to music and time perception through a pilot study that examines participants' affective and temporal feedback on selected Egyptian music excerpts. This interdisciplinary study proposes a cognitive model that explains how temporal meaning arises from music perception and affective experiences.

References

Barsalou, L. W. (2008). *Grounded cognition*. *Annual Review of Psychology, 59*, 617–645.

Droit-Volet, S., Bigand, E., Ramos, D., & Bueno, J. L. O. (2013). Time flies with music whatever its emotional valence. *Acta Psychologica, 143*(1), 11–22. https://doi.org/10.1016/j.actpsy.2013.02.008

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell.

Wharton, T. (2009). *Pragmatics and non-verbal communication*. Cambridge University Press. https://doi.org/10.1017/CBO9780511635649

Equivocality in vocality:

Are emotions perceived like they are expressed?

Maël Maucha

University of Geneva

mael.mauchand@unige.ch

In everyday communication, there is no formal way to guarantee that the emotions we attribute to our interlocutor's vocal signals, such as prosody, truly reflect their internal state. While the correspondence between intended (illocutionary) and effective (perlocutionary) meaning is a recurring challenge in speech, affective prosody lacks the semantic (locutionary) consistency of linguistic signals. Instead, emotional speakers and listeners can only rely on equivocal information, such as subtle acoustic changes, without ever confirming the efficacy of their interaction.

Experimentally, emotion communication research faces a similar issue: emotional representations are usually determined externally, either a priori (by asking participants to produce a determined emotion) or a posteriori (by categorizing vocal stimuli into perceived emotions). Both methods simply displace the problem, and evaluating perceptual responses to such stimuli boils down to comparing two perceptual representations, bypassing the genuine emotion of the speaker. Capturing the true end-to-end, encoding-to-decoding route of emotion communication requires methods considering the emotion felt, the emotion perceived, and the acoustic information mediating the two. Potential paths to close this gap can be found in the presently highlighted research across each step of the lens model of affective speech, combining acoustic, neuropsychological, developmental, and cross-cultural evidence to provide insights into the multifaceted processes facilitating (or hindering) emotional communication. Studies include developmental characterizations of infant emotional vocalizations, multi-dimensional analyses of emotive speech acts, and neuropsychological investigations comparing vocalized, subvocalized, and perceived emotions at the interpersonal level.

Overall, shared emotional and acoustic representations between speaker and listener guarantee some base alignment between encoding and decoding of vocal affective signals. On the speaker's end, the alignment is facilitated by consistency in expressing affective prosodic patterns acquired early in development. On the listener's end, it depends on familiarity with these patterns, both at the cultural and individual level. Acoustically, beyond typical emotion-related acoustic features such as pitch (F0) and loudness, listeners can access and integrate a wide range of other parameters related to voice quality, spectrum, and rhythm to form more fine-grained and complex representations of their interlocutor's affective stance. Often, however, emotion encoding and decoding exhibit partial discrepancies: certain encoded features do not affect perceptual representations, which in turn rely on cues that may not initially reflect the expressed emotion. In the brain, emotional speech representation, expression, and perception seem to involve common motor, pre-motor, and affect-related regions. This suggests forms of embodied emotion and/or empathic processing where voice affect decoding partly mirrors its encoding, with complementary involvement of higher order mentalization networks. Acoustic, cultural, and individual factors in speaker-listener alignment are also reflected across these neural processes.

Together, these findings incorporate into a pragmatic perspective of emotional expression where vocal emotional signals are not mere information capsules but interpersonal acts with an illocutionary force. Emotion communication research, just like everyday emotion communication, must involve efforts to reconcile the emotion felt, the emotion expressed, and the emotion perceived in complex interactive contexts.

"That sounded rude. Were you ironic?"

Emotions and ironic tone of voice

Francesca Panzeri

University of Milan - Bicocca

francesca.panzeri@unimib.it

Ironic remarks are evaluative comments in which speakers manifest their contemptuous attitude towards the thought echoed by the remark (Wilson & Sperber, 2012) or toward the person who would be so foolish as to entertain that thought (Clark & Gerrig, 1984). This judgmental stance implies a strong affective and emotional dimension as well (Gibbs et al., 2002), and, especially in the case of sarcasm, speakers display negative emotions such as sadness, anger, disdain, disgust (Utsumi, 2000; Amenta et al., 2013), even if irony is often associated to jocularity and friendliness (Mauchand et al., 2020).

Our goal was to verify whether the typical intonational profile with which ironic remarks are pronounced encodes these negative emotions. To this aim, we collected a dataset comprising 768 audio tracks of 8 Italian-speaking adults, who uttered 24 statements in both a sincere and an ironic tone. We then used Emozionalmente (Catania et al., 2025) for the automatic detection of emotions transmitted through the vocal channel, to identify the emotions encoded in sincere and ironic remarks of our dataset. We found that ironic remarks display more disgust, more surprise, and more joy, compared to their sincere counterparts.

References

Amenta et al. (2013). Decoding of emotional components in complex communicative situations (irony). *Alcoholism*.

Catania et al. (2025). Emozionalmente. *IEEE Transactions on Audio, Speech and Language Processing*.

Clark & Gerrig (1984). On the pretense theory of irony. *JEP*: General.

Gibbs et al. (2002). What's special about figurative language in emotional communication?. In *The verbal communication of emotions: Interdisciplinary perspectives*.

Mauchand et al. (2020). Irony, prosody, and social impressions of affective stance. *Discourse Processes*.

Utsumi (2000). Verbal irony as implicit display of ironic environment. *Journal of pragmatics*

Wilson & Sperber (2012). Meaning and relevance.

Voicing affect in simultaneous interpreting

Fabrizio Gallai, Pauline Madella & Chara Vlachaki

University IULM, Milan; University of Bedfordshire; University of Brighton

fabrizio.gallai@iulm.it

This paper investigates how prosodic features convey affect in simultaneous interpreting, arguing that affective meaning is a central component of cross-pragmatic expertise, rather than a peripheral by-product.

Drawing on Relevance Theory (Sperber & Wilson 1986/1995), the tool of affective effects (Wharton & de Saussure 2023), and the role of multimodality in interaction (Wharton & Madella 2024) our presentation follows a multimodal analysis of interpreted speech. Within this framework, affective effects are understood as systematic cognitive responses elicited by prosodic and embodied cues. Such cues operate as natural codes, engaging fast evaluative pathways while interacting with slower, propositional reasoning. This approach helps explain why non-verbal behaviours can communicate emotional stance, intensity or tension even where such meanings are difficult to paraphrase and may, at times, exceed speakers' conscious intentions.

This paper explores how prosodic, facial, and gestural cues in the source texts influence interpreters' anticipatory processing, segmentation choices, mitigation or intensification strategies, and vocal alignment. We also explore how interpreters' affective cues guide listeners' attention, trigger emotional evaluation, and shape inferential processing in real time by paying attention to how interpreters reproduce, attenuate or recalibrate affective information through their own prosodic choices. The dataset consists of a multilingual corpus from a European Parliament debate, comprising nine speeches in four languages (French, Greek, Italian and Portuguese) and their English renditions. The paper contributes to interdisciplinary discussions within relevance-theoretic pragmatics and affective science by highlighting the role of prosody in the communication of affect.

References

Sperber, Dan, Wilson, Deirdre, 1986/1995, *Relevance: Communication and Cognition*, Harvard U.P., Cambridge (MA).

Wharton, Tim, De Saussure, Louis, 2023, *Pragmatics and Emotion*, C.U.P., Cambridge.

Wharton, Tim, Madella, Pauline, 2024, "Non-Verbal Communication and Context: Multimodality in Interaction", in J. Romero-Trillo (ed.), *The Cambridge Handbook of Language in Context*, C.U.P., Cambridge, pp. 419-36.

Can affective prosody contribute to the derivation of pragmatic meanings?

Evidence from masked speech

Nikos Vergis

Hellenic Open University

vergis.nikolaos@ac.eap.gr

Prosody has been found to contribute significantly to the computation of pragmatic meanings such as speech acts and im/politeness, yet the mechanisms through which this is instantiated are debated. Some have argued that pragmatic meanings arise only through a combinatorial computation of contextual factors, while others have provided evidence that there are unique prosodic signatures that map onto specific pragmatic meanings. The present study puts forward the hypothesis that the derivation of pragmatic meanings is partly possible due to the inherent power of prosody to convey affect and affective states. In particular, I investigated im/politeness meanings by teasing apart the effect of prosody from the lexical content of the utterance. Two types of stimuli selected from a validated database of im/polite utterances were employed: unfiltered utterances (lexical content + prosody) and masked versions of the same utterances (in which only prosody was retained through low-pass filtering, and semantics was removed). Listeners (N=40) were asked to rate both types of stimuli on several scales: two scales of affectivity (valence and arousal) (Exp. 1) and a scale of im/politeness (Exp. 2). Results showed that both unfiltered and masked conditions exhibited sensitivity to valence, revealing that rude utterances were perceived as more negative than polite ones (Exp. 1). Also, listeners were able to differentiate levels of im/politeness in rude and polite utterances even when semantics was removed (Exp. 2). Multiple regressions also showed that valence predicted im/politeness, and that both gradient and categorical acoustic measures of our stimuli accounted for variance in vocal im/politeness. These results highlight the role of affective prosody in pragmatic inference and contribute to a wider discussion in the context of inferentialist and non-inferentialist approaches to meaning.

I listened to *His Dark Materials* and was transported to another world:

Audiobooks as emotional portals

Chara Vlachaki

University of Brighton

c.vlachaki2@brighton.ac.uk

This presentation explores the ways audiobooks impact the listener's affective states using examples from Philip Pullman's trilogy *His Dark Materials*. I claim that prosodic features through the narrator performance function as emotional cues and shape the listener's experience: in effect, the audiobook becomes an emotional portal to a vivid, immersive 'other world'. Audiobooks, hybrid, multimodal aesthetic objects are a place where audio and textual elements converge. This interplay between audio performance and text has been linked to physiological evidence of increased heart rate indicating increased cognitive and affective engagement (Richardson et al. 2018). From the perspective of relevance theory (Sperber & Wilson 1995), cognitive engagement can be understood in terms of a balanced relationship between cognitive effects and processing effort. However, the interpretation of literature often requires more effort. This can be addressed in the form of a cognitive/affective loop (Wharton & de Saussure 2023). In this presentation, I will discuss this loop by exploring interpretive resemblance in audiobooks. I argue that the narrator's performance interpretively resembles the written text aiming to convey as closely as possible what the author intended to communicate thus, guiding the audience to construct mental representations triggering an array of propositions as well as emotions and feelings. The latter has been defined as a loose use of translation (Vlachaki 2025).

References

Richardson, D., Griffin, N., Zaki, L., Stephenson, A., Yan, J., Curry, T., Noble, R., J., Hogan, J., Skipper, J. & Devlin, J. (2018) Measuring narrative engagement: The heart tells the story. [preprint] https://doi.org/10.1101/351148

Sperber, D. & Wilson, D. (1986/1995) *Relevance: Communication and cognition*. Blackwell.

Vlachaki, C. (2025) 'Translating the ineffable in art', in F. Alves & F. Gallai (eds.) *Applications of relevance theory to translation and interpreting*. Routledge, pp. 222-242.

Wharton, T. & Louis, de Saussure (2023) *Pragmatics and emotion*. CUP.