

The Neurolinguistics of Attitudinal Prosody

Chidera Elvis Mbah

University of Essex

Abstract

If prosody is the technical term for tone of voice, attitudinal prosody should refer to how the tone of voice communicates information about the speaker's attitude. Yet it is still unclear whether attitudes are sufficiently distinct from emotions to be communicated through prosodic cues. This essay will suggest that attitudinal prosody can indeed be distinguished from emotional prosody. To this end, the literature concerning the conceptual distinctions between attitudes and emotions will be reviewed. The essay will then summarise recent neurocognitive evidence from clinical studies, neuroimaging, and electroencephalography concerning the correlates of attitudinal prosody. It will conclude with suggestions for further study, given the relatively scant research in this area.

Keywords: attitudes, prosody, neurolinguistics, functional neuroimaging, electroencephalography

Introduction

Social harmony greatly depends on how well people communicate with one another. Successful communication, in turn, depends on the ability to understand both verbal and nonverbal cues. Though a good amount of information can be transmitted through the sheer meaning of words, subtler and arguably more crucial information can be communicated through nonverbal elements. In particular, the tone of voice contributes contextual information that helps a listener discern not just the speaker's meaning but the speaker's intentions as well (Cole, 2015; Wilson and Wharton, 2006). The technical term for tone of voice is "prosody", and it has been described as referring to "suprasegmental aspects of speech that are paralinguistic" (Mitchell and Ross, 2013, p. 471). Behrens (1985, p. 332) called it "the melody of speech", indicating through this metaphor that prosody encompasses the more "musical" aspects of spoken language, such as rhythm, stress, intonation, amplitude, tempo, timbre, fundamental frequency, and, perhaps most importantly, pitch. Prosody

The neurolinguistics of attitudinal prosody

is an essential cue for grasping the meaning of speech, since most language is ambiguous, admitting of more than one meaning, and humans may have evolved to let out their true feelings and intentions in this way (Turner, 2002). Monrad-Krohn (1947) specified four categories of prosody. *Intrinsic prosody* refers to prosody that identifies the linguistic characteristics of speech, such as the distinction between a question and a statement. *Intellectual prosody* concerns the speaker's intellectual stance towards the information being conveyed. This category encompasses what may be called *attitudinal prosody*, which describes how prosodic cues reveal people's beliefs and attitudes. *Emotional prosody* indicates how emotions are communicated by tone of voice. Finally, *inarticulate prosody* refers to sub-linguistic sounds which, though nonverbal, add contextual information to discourse (including such exclamations as "oh!", "ah", etc.). This essay aims to examine the neurolinguistics of attitudinal prosody. Specifically, this essay will review linguistic and neurocognitive evidence for attitudinal prosody as an important element of human communication.

Attitudinal Prosody: Is There Such a Thing?

Attitudinal prosody has often been classified with emotional prosody under the title of *affective prosody*, both being treated as indistinguishable from each other (Blanc and Dominey, 2003). Indeed, these forms of prosody share certain features and possibly recruit similar neural pathways (Wickens and Perry, 2015). This is not surprising because emotions and attitudes are intertwined with each other. Yet there seem to be important conceptual differences between them. Attitudes, for one, are the dispositions of a person towards a thing, another person or group of people, an event, or some other phenomenon (Mitchell and Ross, 2013). Emotions, however, have been defined as mental experiences "with high intensity and high hedonic content" (Cabanac, 2002, p. 69). In other words, attitudes refer to a person's intellectual "stance" or "posture" towards a subject or object, whereas emotions reflect a person's affective state. Now, this distinction is not so clear-cut, because attitudes often go together with emotions, and vice versa. But it would not be accurate to speak of irony, for example, as an emotion; nor of doubt, deception, or confidence. The ways of being described by these concepts are intentional; that is, they reflect a person's attitudes towards others. Thus, one could claim that while emotions refer to feelings, attitudes concern beliefs or behaviours (Wickens and Perry, 2015). For example, Mary is *being* sarcastic towards Lisa because she *feels* angry with her. Attitudes are also more enduring than emotions. This is probably because

they are indicative of entrenched beliefs, whereas emotions are somewhat fleeting, often being driven by specific situations (Scherer, 2003). Furthermore, theorists have pointed out that emotions tend to be the involuntary expressions of more deliberate and socially-controlled attitudes (Aubergé, 2002).

In addition to these conceptual distinctions, although attitudinal and emotional prosody both rely on a limited range of vocal cues, particularly pitch, emotional prosody seems also to depend more on timbre (i.e., harshness or softness) and attitudinal prosody on “prosodic contour” (i.e., the modulation of rhythm; Scott *et al.*, 2010). Other particularly important acoustic indicators of attitudes include amplitude and loudness (Mitchell and Ross, 2013). Confident speech, for example, is characterised by loudness and a steady rhythm. By contrast, a lack of confidence betrays itself by halting pace and heightened pitch (Monetta *et al.*, 2008). Similarly, rising and falling intonations are associated with agreement and disagreement respectively (Gussenhoven, 2004; Juslin and Laukka, 2003). Whether prosody conveys an attitude or an emotion sometimes depends on the context. For example, slight changes in stress and fundamental frequency, in addition to contextual clues, can indicate that the speaker is either surprised or impressed (Cruttenden, 1986). A mismatch between these vocal features and the context sometimes creates unintended effects. Thus, attitudinal prosody may create a “diffuse impression of friendliness, condescension, nervousness, etc., which may never surface to consciousness” (Wilson and Wharton, 2006, p. 1565).

Attitudes commonly communicated by prosody include curiosity, disagreement, puzzlement, incredulity, suspicion, and belief (Morlec *et al.*, 2001). Speakers may modulate vocal rhythm, intonation, or pitch to indicate that they are speaking the truth and should be trusted. They may also evince vocal signals of affiliation, or demonstrate respect towards their interlocutors (Pell, 2006). Some researchers, however, have claimed that prosodic features indicate only affect but not specific attitudes (e.g., Mozziconacci, 2001). Though this stance is understandable, given the difficulty in disentangling emotions from attitudes, it does not seem to account for many aspects of human social communication. Moreover, prosody has been identified as an important indicator of attitudes in diverse social situations (Jiang and Pell, 2017; Mauchand *et al.*, 2020).

Concerning ironic attitudes, specifically, Bryant and Fox Tree (2005) argued against there being such a thing as an ironic tone of voice. Ironic attitudes are communicated through a deliberate

mismatch between the surface meaning of an utterance and its underlying or real meaning (e.g., sarcastic comments). Bryant and Fox Tree suggested that prosodic features become important for recognising irony only when other contextual information is missing. For hostile irony (or sarcasm), however, significant prosodic effects have been pinpointed. In the English language, it appears that, relative to other expressions of attitudes, sarcastic statements are characterised by decreased fundamental frequency, decreased variation in frequency, coarser timbre, and an attenuated pace (Cheang and Pell, 2008). Besides, Macuchand *et al.* (2020) examined irony comprehension by presenting listeners with compliments and criticisms and asking them to rate the attitude of the speaker. The researchers discovered an asymmetry in the perception of ironic attitudes compared to literal attitudes, such that people judged ironic compliments (or teasing) as less friendly than literal compliments, and ironic criticisms as less hostile than literal criticisms. That the contrast between both attitudes was communicated solely by prosodic variations suggests that prosody does allow listeners to distinguish between different kinds of irony.

The Neuroanatomy of Attitudinal Prosody

Given the degree of overlap between emotional and attitudinal prosody, it is reasonable to suspect that they recruit similar brain structures and networks (Mitchell and Ross, 2013). The influential multi-stage model of emotional language processing identifies three broad processing stages (Kotz and Paulmann, 2011). In the first stage, around 100 milliseconds (ms) after the onset of vocal input, initial acoustic properties (e.g., pitch) are extracted and analysed; this stage involves the activity of bilateral auditory cortices (Kotz and Paulmann, 2011). During the second stage, vocal input can be distinguished in terms of its affective relevance. Indeed, it has been shown that as early as 200 ms post-acoustic onset, emotional prosodies are discernible from one another and from neutral ones (Paulmann, 2016). This second stage is thought to recruit the right anterior superior temporal gyrus and sulcus (Jiang and Pell, 2015). The final stage of processing, in contrast to the earlier ones, involves “higher-level” cognition, particularly the integration of contextual information with other sources (e.g., prosodic, semantic, and syntactic cues). This late stage, which occurs around 400 ms after the presentation of auditory input, primarily involves the bilateral inferior and orbitofrontal cortex (Paulmann, 2016). Mitchell and Ross (2013) speculated that this late stage is when the processing of attitudes deviates from emotional processing. They also suggested that the two early stages are more or less similar no matter what kind of prosody is involved; it is only at the latter

stage that specific emotions and attitudes can be identified. This would imply a close similarity, if not actual correspondence, of attitudes with complex emotions that require higher cognitive processes to disentangle (Wickens and Perry, 2015). Further research would be required to tease out the differences between attitudes and emotions in this late processing stage.

A few studies have directly investigated attitudinal prosody in clinical populations. These generally suggest a strong role of the right hemisphere. For example, patients with a lesion in their right hemisphere (RH) have shown difficulty in inferring politeness using politeness or confidence cues from sentences (Pell, 2006). Pell (2007) also studied the processing of speaker attitudes from a group of patients with RH lesions. The patients were presented with two kinds of stimuli. In the linguistic task, the patients were presented with spoken sentences that provided clear lexical signals of confidence (e.g., *surely*, *perhaps*, *most likely*). In the prosody task, they heard pseudo-sentences that portrayed various degrees of confidence by modulating prosody. The results showed that the RH patients were comparatively worse than healthy controls (HC) at distinguishing degrees of confidence from prosodic cues (Pell, 2007). In another experiment, RH patients heard sentences that adjusted different levels of politeness by using either lexical-semantic cues or prosodic features alone. Compared to HC, the patients with RH damage were significantly worse at recognising the speaker's level of politeness (Pell, 2007). These findings, however, should be interpreted with caution due to the study's relatively low sample size. Moreover, only healthy controls were used in Pell's study; further research will need to use patients with other patterns of damage to demonstrate a double dissociation. Patients with epilepsy in the right posterior superior lateral temporal cortex have also shown significant impairment in their ability to identify speaker attitudes from prosody (Tompkins and Mateer, 1985). These studies suggest that the processing of attitudes, like that of emotions, may be mediated by the right hemisphere. Other brain areas seem to serve important functions in the detection of attitudes from speech. Basal ganglia, for example, have been linked to the processing of attitudinal prosody. Monetta *et al* (2008) studied attitude detection in people with Parkinson's Disease (PD), a condition characterised by basal ganglia dysfunction. The patients listened to pseudo-utterances and rated the confidence and politeness levels of speakers. The PD patients were markedly less able than HC to differentiate between different levels of confidence. For politeness, PD patients performed almost comparably with HCs; however, they were less sensitive to low politeness (Monetta *et al.*, 2008). Thus, damage to the basal ganglia seemed to impair

the accurate detection of attitudes from prosodic cues—at least, for such attitudes as confidence and politeness.

As Mitchell and Ross (2013) cautioned, the complex nature of attitudes makes it difficult to identify the functional neuroanatomy of attitudinal prosody from imaging studies. They hypothesised that any differences in the brain regions mediating emotional and attitudinal prosody are likely to be found in the frontal lobe. This hypothesis is in keeping with the results of their meta-analysis of neuroimaging studies, in which they found evidence for a bilateral temporofrontal network governing the detection of emotional prosody (Mitchell and Ross, 2013). Unfortunately, there have been virtually no imaging studies dealing directly with attitudinal prosody. This is a gap in the literature that needs to be filled.

Similarly, there have been few attempts to investigate the link between prosody and attitude detection using electroencephalography (EEG). As briefly mentioned before, event-related potentials (ERPs) for various stages of emotional processing have been identified. Notably, the N100 signal has been interpreted as an index of early auditory sensory processing (Jiang and Pell, 2015; Paulmann and Kotz, 2008). The fronto-central P200 corresponds with emotion detection and has been known to distinguish emotion-laden speech from neutral utterances (Paulmann, 2016). There is evidence that the P200 signal amplitude differs between various kinds of emotion (Paulmann *et al.*, 2013). Apart from these relatively early signals, a late positivity effect (typically P600) has been associated with higher cognitive processing in various contexts (Kolk and Chwilla, 2007; Regel *et al.*, 2014). According to Mitchell and Ross (2013), this late positivity signal should serve to distinguish attitudinal prosody from other kinds of prosody.

In this vein, Wickens and Perry (2015) sought to measure using EEG the differences between emotional prosody (anger) and attitudinal prosody (sarcasm). Anger and sarcasm are closely related, the latter frequently being a behavioural manifestation of the former. Wickens and Perry adopted the “prosodic expectancy violation” paradigm (Paulmann and Kotz, 2008), in which participants heard cross-spliced “syntactically matching and semantically neutral” sentences with a neutral beginning (*He has*) and either a neutral, an angry, or a sarcastic prosodic ending (*a serious face*; Wickens and Perry, 2015, p. 5). They compared the effect of task design by performing two experiments with different experimental tasks. In the first experiment, the participants were tasked

to identify which attitude or emotion each sentence was conveying. There was, however, no significant difference in the processing of angry and sarcastic prosody in the early or late stages, although both differed from neutral prosody. In particular, both elicited an anterior positivity around 200 to 350 ms post-violation. On the other hand, a probe-verification task was used in the second experiment. Here, the participants were required to indicate whether a word displayed on a screen was present in the spoken sentence. The results of this experiment showed a difference in processing between angry and sarcastic prosody within the early negativity window (100 to 200 ms), with a greater negative amplitude for angry prosody. These results suggest that task differences may play a role in the processing of attitudinal prosody, but further studies are needed to make any stronger conclusions.

Suggestions for Future Study

The foregoing has given some indication that attitudinal prosody, while interesting for its implications in human social communication, has hardly been systematically investigated. Further discoveries may prove that the differences between prosodies for emotion and attitudes are so slight that they cannot really be separated. But it will be premature to draw this conclusion until the topic has been deliberately studied. In particular, further attention should be dedicated to the conceptual definitions of attitudes and emotions, including the psychoacoustic overlap and differences between them. Furthermore, researchers should study the impairments in the perception of attitudes across clinical populations, given the paucity of research in this area. In addition, imaging studies (both fMRI and EEG) should specifically target attitudinal prosody perception in both clinical and healthy populations. Finally, as research confirms a substantial distinction between attitudinal and emotional prosody, efforts should go into developing corpora for attitudinal prosody that is comparable in range, size, and validity with those for emotional prosody (Mitchell and Ross, 2013).

Conclusion

This essay has discussed attitudinal prosody in an attempt to show that it is distinct from emotional prosody and worthy of study in its own right. To this end, I reviewed a body of research concerning the conceptual and theoretical definitions of attitudes, the psycholinguistic delineations of prosodic cues for attitude, the brain structures and regions involved in the processing of attitudinal prosody, and the electrical activity associated with that processing. The difficulty of clearly distinguishing

between prosodies for attitudes and those for emotions suggests the need for further research into this important area.

References

- Adolphs, R. (2002) 'Neural systems for recognizing emotion', *Current Opinion in Neurobiology*, 12(2), pp. 169-177. doi: [https://doi.org/10.1016/s0959-4388\(02\)00301-x](https://doi.org/10.1016/s0959-4388(02)00301-x)
- Behrens, S. J. (1985) 'The perception of stress and lateralization of prosody', *Brain and Language*, 26(2), pp. 332-348. doi: [https://doi.org/10.1016/0093-934X\(85\)90047-1](https://doi.org/10.1016/0093-934X(85)90047-1)
- Blanc, J. M. and Dominey, P. F. (2003) 'Identification of prosodic attitudes by a temporal recurrent network', *Cognitive Brain Research*, 17(1), pp. 693-699. doi: [https://doi.org/10.1016/s0926-6410\(03\)00195-2](https://doi.org/10.1016/s0926-6410(03)00195-2)
- Bosco, F. M., Parola, A., Valentini, M. C., and Morese, R. (2017) 'Neural correlates underlying the comprehension of deceitful and ironic communicative intentions', *Cortex*, 94(1), pp. 73-86. doi: <https://doi.org/10.1016/j.cortex.2017.06.010>
- Bryant, G. A. and Fox Tree, J. E. (2005) 'Is there an ironic tone of voice?' *Language and Speech*, 48(3), pp. 257-277. doi: <https://doi.org/10.1177/00238309050480030101>
- Cabanac, M. (2002) What is emotion? *Behavioural Processes*, 60(2), pp.69-83. doi: [https://doi.org/10.1016/S0376-6357\(02\)00078-5](https://doi.org/10.1016/S0376-6357(02)00078-5)
- Cheang, H. S. and Pell, M. D. (2008) 'The sound of sarcasm', *Speech Communication*, 50(5), pp. 366-381. <https://doi.org/10.1016/j.specom.2007.11.003>
- Cole, J. (2015) Prosody in context: A review. *Language, Cognition and Neuroscience*, 30(1-2), pp.1-31. doi: <https://doi.org/10.1080/23273798.2014.963130>
- Cruttenden, A. (1997) *Intonation*. Cambridge University Press.
- Grichkovtsova I., Morel M., and Lacheret, A. (2012) 'The role of voice quality and prosodic contour in affective speech perception', *Speech Communication*, 54(1), pp. 414-429.
- Gussenhoven, C. (2004) *The phonology of tone and intonation*. Cambridge University Press.

Jiang, X. and Pell, M. D. (2015) 'On how the brain decodes vocal cues about speaker confidence', *Cortex*, 66(1), pp. 9-34. doi: <https://doi.org/10.1016/j.cortex.2015.02.002>

Jiang, X. and Pell, M. D. (2017) 'The sound of confidence and doubt', *Speech Communication*, 88(1), pp. 106-126. doi: <http://dx.doi.org/10.1016/j.specom.2017.01.011>

Juslin, P. N. and Laukka, P. (2003) 'Communication of emotions in vocal expression and music performance: Different channels, same code?', *Psychological Bulletin*, 129(5), pp. 770-814. doi: <https://doi.org/10.1037/0033-2909.129.5.770>

Kolk, H. and Chwilla, D. (2007) 'Late positivities in unusual situations', *Brain and Language*, 100(3), pp. 257-261. doi: <https://doi.org/10.1016/j.bandl.2006.07.006>

Kotz, S. A. and Paulmann, S. (2011) 'Emotion, language, and the brain', *Language and Linguistics Compass*, 5(3), pp. 108-125. doi: <https://doi.org/10.1111/j.1749-818X.2010.00267.x>

Mauchand, M., Vergis, N., and Pell, M. D. (2020) 'Irony, prosody, and social impressions of affective stance', *Discourse Processes*, 57(2), pp. 141-157. doi: <https://doi.org/10.1080/0163853X.2019.1581588>

Mitchell, R. L. C. and Ross, E. D. (2013) 'Attitudinal prosody: What we know and directions for future study', *Neuroscience and Biobehavioural Reviews*, 37(3), pp. 471-479. doi: <https://doi.org/10.1016/j.neubiorev.2013.01.027>

Monrad-Krohn, G. H. (1947) 'Dysprosody or altered melody of language', *Brain: A Journal of neurology*, 70(1), pp. 450-415. doi: <https://psycnet.apa.org/doi/10.1093/brain/70.4.405>

Morlec, Y., Bailly, G., and Aubergé, V. (2001) 'Generating prosodic attitudes in French: data, model and evaluation', *Speech Communication*, 33(4), pp. 357-371. doi: [https://doi.org/10.1016/S0167-6393\(00\)00065-0](https://doi.org/10.1016/S0167-6393(00)00065-0)

Mozziconacci, S. J. (2001) 'Modeling emotion and attitude in speech by means of perceptually based parameter values', *User Modeling and User-Adapted Interaction*, 11(4), pp. 297-326. doi: <https://doi.org/10.1023/A:1011800417621>

Paulmann, S. (2016) 'The neurocognition of prosody' in *Neurobiology of Language*. Academic Press, pp. 1109-1120. doi: <https://doi.org/10.1016/B978-0-12-407794-2.00088-2>

Paulmann, S., Bleichner, M. and Kotz, S. A. (2013) 'Valence, arousal, and task effects in emotional prosody processing', *Frontiers in Psychology*, 4(2013) [online]. doi:

<https://doi.org/10.3389/fpsyg.2013.00345>

Pell, M. D. (2006) 'Judging emotion and attitudes from prosody following brain damage', *Progress in Brain Research*, 156, pp. 303–317. doi: [https://doi.org/10.1016/s0079-6123\(06\)56017-0](https://doi.org/10.1016/s0079-6123(06)56017-0)

Pell, M.D. (2007) 'Reduced sensitivity to prosodic attitudes in adults with focal right hemisphere brain damage', *Brain and Language*, 101(1), pp. 64–79. doi:

<https://doi.org/10.1016/j.bandl.2006.10.003>

Regel, S., Meyer, L. and Gunter, T. C. (2014) 'Distinguishing neurocognitive processes reflected by P600 effects: Evidence from ERPs and neural oscillations', *PLoS One*, 9(5), e96840 [online]. doi:

<https://doi.org/10.1371/journal.pone.0096840>

Scherer, K. R. (2003) 'Introduction: Cognitive components of emotion' in R. J. Davidson, K. R. Scherer and H. H. Goldsmith (eds.) *Handbook of affective sciences*. Oxford University Press, pp. 563–571.

Scott, S. K., Sauter, D. A., Eisner, F. and Calder, A. J. (2010) 'Perceptual cues in nonverbal vocal expressions of emotion', *Quarterly Journal of Experimental Psychology*, 63, 2251–2272 [online].

doi: <https://doi.org/10.1080/17470211003721642>

Tompkins, C. A. and Mateer, C. A. (1985) 'Right hemisphere appreciation of prosodic and linguistic indications of implicit attitude', *Brain and Language*, 24(2), pp. 185–203. doi:

[https://doi.org/10.1016/0093-934X\(85\)90130-0](https://doi.org/10.1016/0093-934X(85)90130-0)

Turner, H. (2002) 'An introduction to methods for simulating the evolution of language' in A. Cangelosi and D. Parisi (eds.) *Simulating the evolution of language*. London: Springer, pp. 29–50.

doi: https://doi.org/10.1007/978-1-4471-0663-0_2

Wickens, S. and Perry, C. (2015) 'What Do You Mean by That?! An electrophysiological study of emotional and attitudinal prosody', *PLoS ONE*, 10(7): e0132947 [online]. doi:

<https://doi.org/10.1371/journal.pone.0132947>

Wildgruber, D., Ackermann, H., Kreifelts, B. and Ethofer, T. (2006) 'Cerebral processing of linguistic and emotional prosody: fMRI studies', *Progress in Brain Research*, 156, pp. 249–268.

The neurolinguistics of attitudinal prosody

Wilson, D. and Wharton, T. (2006) 'Relevance and prosody', *Journal of Pragmatics*, 38(10), pp. 1559-1579. doi: <https://doi.org/10.1016/j.pragma.2005.04.012>

© Chidera Elvis Mbah. This article is licensed under a Creative Commons Attribution 4.0 International Licence (CC BY).