Research News

Emotional prosody: sex differences in sensitivity to speech melody

Mireille Besson, Cyrille Magne and Daniele Schön

Emotional prosody is defined as the ability to express emotions through variations of different parameters of the human speech, such as pitch contour, intensity and duration. A recent study has discovered clear differences between men and women in the time course of emotional prosodic processing. There are several interpretations of these intriguing results that hold promise for future studies on language comprehension.

Emotional prosody is probably one of the most basic features of language. An example is an infant preferring his or her mother's voice to a stranger's, which has been observed as early as 2–4 days after birth [1]. However, the study of prosody is very complex because, at the same time, it is both universal across human languages and specific to each one [2].

The influence of prosody

Most prosodic systems offer distinctive intonation patterns for questions and declarations (rising versus lowering of pitch contour at the end of the utterance). However, when speaking a second language our accent clearly illustrates the need for different prosodic systems. Moreover, prosody allows communication of both linguistic and emotional intentions at the same time and carries important information related to the sex, age and emotional state of the speaker [3]. Research into the encoding of emotional states in speech signals show clear correlation with global properties, such as loudness, speech rate and pitch contour [4,5]. For example, depressed and schizophrenic patients typically show reduced emotional prosody expression [6], and depressed children (9–11 years old) show less ability than non-depressed children to accurately identify affective prosody [7]. Recently, efforts have been made to explore the specific acoustic features of emotional speech using objective acoustic measures, such as the harmonics-to-noise ratio and the event-related brain potentials (ERPs) method [8,9]. The main advantage of using ERPs is the ability to precisely define the time course of information processing. Moreover, it provides a wonderful tool for studying perceptive and cognitive processes in language comprehension as they unfold in real time.

Exploring sex differences

Using the ERP method, Schirmer *et al.* have studied the time course of the relationship between emotional prosody and word recognition [10]. The novelty in their approach is their comparison between men and women. In an elegant cross-modal priming experiment, the participants listened to semantically neutral German sentences (e.g. 'Yesterday she had her final exam') spoken with either a happy or sad intonation. A semantically related target word with positive or negative valence ('success' or 'failure', respectively) was presented visually after each spoken sentence. The target words either matched or did not match with the emotional sentence prosody (e.g. 'success' or 'failure' following the sentence: 'yesterday she had her final exam', spoken with a happy intonation). In addition, participants performed a lexical decision task (LDT) while their EEG was recorded from 58 scalp electrodes.

Interestingly, the results demonstrate significant differences between men and women in the processing of emotional congruence. When the time interval between the spoken sentences and the visual targets is short (200 ms), women respond faster to matching targets, than to targets that do not match the sentence prosody. Moreover, the N400 component of the ERPs, which is known to reflect word expectancy, is smaller for targets that match compared with those that do not match the sentence prosody. The results indicate that women base their linguistic expectations on emotional prosody as early as 150 ms following visual target onset. By contrast, men do not show any electrophysiological priming effect, but respond faster to positive target words than negative target words, indicating that men process word meaning independently of sentence emotional prosody (Fig 1a). From these results, a possible conclusion could be that men are impermeable to emotional prosody. However, results of a second experiment enabled Schirmer *et al.* to establish that this is not the case; men are sensitive to emotional prosody, but they are slower than women to process it (Fig. 1b).

The results highlight a puzzling dissociation between men and women in the processing of emotional prosody, but the reasons for it are not obvious. In previous studies, results have shown stronger left lateralization of language functions for men compared with women [11]. The dichotic listening test used by Schirmer *et al.* did not, however, reveal any sex difference in language lateralization. Interestingly, results of a recent experiment by Nowicka and Fersten [12] have revealed longer left-to-right compared with right-to-left interhemispheric transmission times (IHTT) for men, whereas both directions are similar in women.

The generally accepted hypothesis is that emotional prosody processing is lateralized to the right hemisphere [13], whereas lexical processing is left lateralized. Therefore, the longer onset-latency of the emotional prosodic priming effect observed in men might result from longer left-to-right IHTT for men. However, it should be noted that such an audacious interpretation requires further support because shorter IHTT in men compared with women has also been reported [14]. Clearly, these discrepancies might result from the task used and the kind of stimuli employed [15].

Directions for future research

The results reported by Schirmer *et al.* highlight intriguing possibilities that need to be tested further. For instance, it is important to determine the extent to which these results are task-dependent. Therefore, the results from a LDT (that requires access to word meaning) should be compared with results from a task requiring prosodic judgement (i.e. whether the

http://tics.trends.com 1364-6613/02/$ – see front matter PII: S1364-6613(02)01975-7

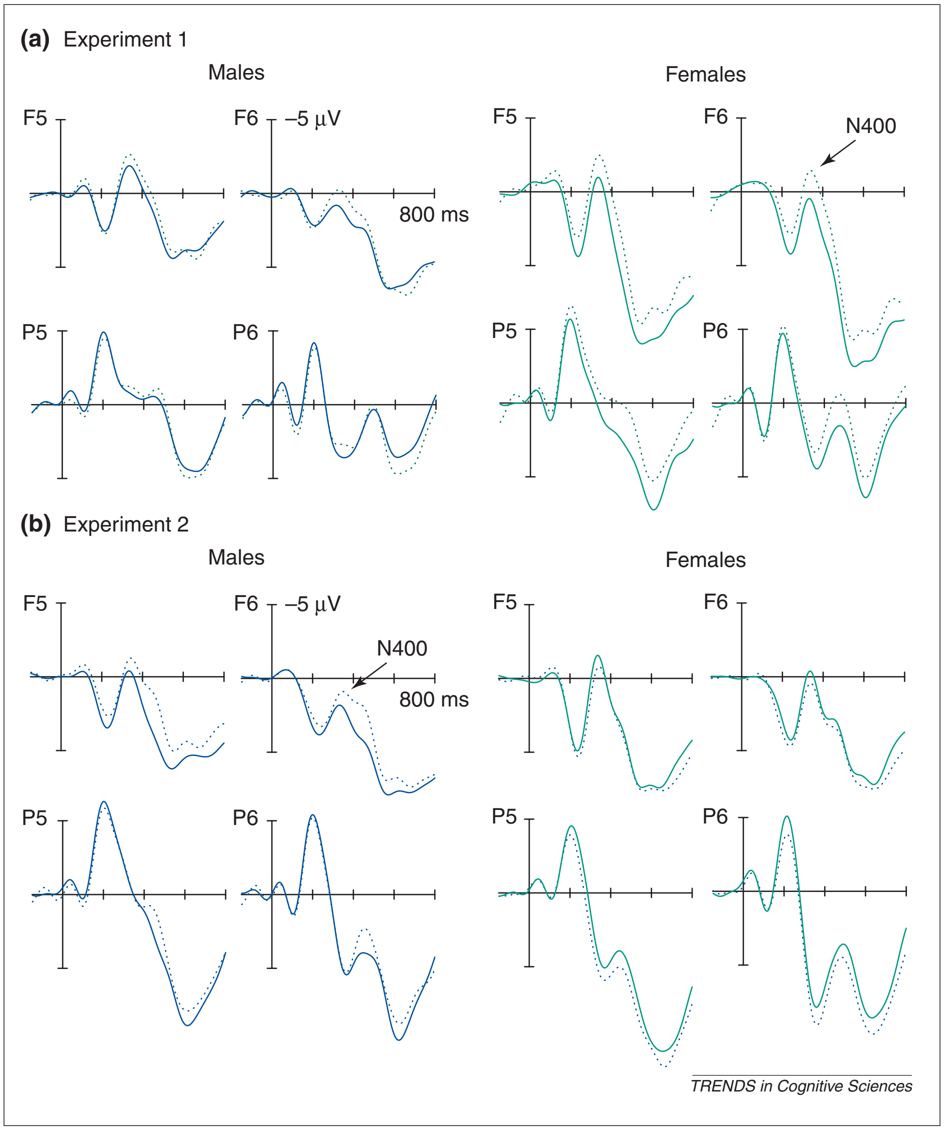

Fig. 1. Time course of emotional prosodic processing in men and women [10]. The time interval (T) between spoken sentences (primes) and subsequently presented visual target words were different in the two experiments. (a) T = 200 ms; women (right plots) clearly show an emotional priming effect in the ERPs (faster response to matching targets), seen from 150 ms following target word onset and culminating in the N400. Solid lines represent matches between primes and targets; broken lines represent mismatches. No such effect is found in men (left). (b) T = 700 ms; the reverse pattern of results is observed.

sentence is spoken with a happy or a sad intonation). Manipulating the direction of attention should enable determination of the extent that prosodic processing relies on automatic or controlled processing. We have attempted this in our laboratory by independently manipulating semantic and prosodic congruency. Participants were asked to focus their attention on either the semantics or on the prosody of the spoken sentences. The results show that although participants could focus on the semantics of the sentence and ignore the prosody, they could not ignore the semantics while focussing on the prosody.

In addition, it is important to determine whether the results reported by Schirmer *et al.* are stimulus-dependent. As detailed elsewhere [16], music is often considered the language of emotion. Therefore, it is interesting to compare emotional prosody with music, and determine whether the parameters that convey emotion in music and language are similar or different. We recently conducted an experiment aimed at directly comparing pitch-contour processing in music and language. Briefly, results have shown that pitch-contour violations elicit similar patterns of brain waves in music and in language. Musicians were compared with non-musicians, and both groups comprised men and women.

In light of the Schirmer *et al.* results, it would be interesting to compare men with women using a pitch violation detection task. In addition, comparing male musicians and male non-musicians in the cross-modal priming experiment used by Schirmer *et al.*, should help determine whether male musicians show faster emotional priming effects than male non-musicians.

Clearly, the study of the prosodic aspects of language processing is a very rich domain, and is just beginning to be explored using brain-imaging methods. Aside from the ERP studies discussed here, interesting results are emerging from experiments using functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (TEP).

Using fMRI, Buchanan *et al.* [17] have compared emotional prosodic processing and phonemic processing. Participants were asked to discriminate between words based on either emotional prosody (angry, happy, sad or neutral) or phonemic characteristics (words starting with different consonants e.g. bower, power, tower and so on). Results show that the detection of emotional prosody is associated with increased activation in the right inferior frontal lobe and right anterior auditory cortex, whereas detection of phonemic differences elicit larger activation in the left inferior frontal lobe. Thus, processing the linguistic and emotional parameters of speech signals might involve specific neural networks. Gandour *et al.* have reached a similar conclusion based on cross-linguistic PET studies [18].

Concluding remarks

Results from the study by Schirmer *et al.* are relevant to the fundamental issues in language research regarding the use of non-linguistic information (emotional prosody) in language comprehension and production. The idea that the human perceptive and cognitive systems make use of all the available information in the incoming speech signal (i.e. acoustic, phonemic, morphemic, emotional, semantic, syntactic and pragmatic) to process meaning and affective meaning, needs to be addressed using different approaches. Among the most exciting challenges is to determine whether these different aspects of the speech signal are processed (1) sequentially or in interaction, and (2) by specific, dedicated modules or by general perceptive and cognitive mechanisms. Although these are difficult issues, they must be researched to improve our understanding of the relationship between brain and behaviour. There are exciting implications for the rehabilitation of patients with language disorders, and for man–machine communication in the real and virtual world.

Acknowledgements

This paper was written while Mireille Besson benefited from a grant from the International Foundation for Music Research (IFRM: RA #194).

References

1 Hepper, P.G. *et al.* (1993) Newborn and fetal response to maternal voice. *J. Reprod. Infant Psychol.* 11, 147–153
2 Lacheret-Dujour, A. and Beaugendre, F., eds (1999) *La prosodie du français*. CNRS Editions
3 Johnstone, T. and Scherer, K.R. (2000) Vocal communication of emotion. In *Handbook of Emotions* (Lewis, M. and Haviland-Jones, J.M., eds), pp. 220–235, Guilford Press
4 Ladd, D. *et al.* (1985) Evidence for the independent function of intonation contour type, voice quality, and F0 range in signalling speaker affect. *J. Acoust. Soc. Am.* 78, 435–444
5 Scherer, K.R. *et al.* (1991) Vocal cues in emotion encoding and decoding. *Motiv. Emotion.* 15, 123–148
6 Alpert, M. *et al.* (2001) Reflections of depression in acoustic measures of the patient's speech. *J. Affect. Disord.* 66, 59–69
7 Emerson, C.S. *et al.* (1999) Investigation of receptive affective prosodic ability in school-aged boys with and without depression. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 12, 102–109
8 Alter, K. *et al.* Affective encoding in the speech signal and in event-related brain potentials. *Speech Commun.* (in press)
9 Steinhauer, K. *et al.* (1999) Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nat. Neurosci.* 2, 191–196
10 Schirmer, A. *et al.* (2002) Sex differentiates the role of emotional prosody during word processing. *Cogn. Brain Res.* 14, 228–233
11 Shaywitz, B.A. *et al.* (1995) Sex differences in the functional organization of the brain for language. *Nature* 373, 607–609
12 Nowicka, A. and Fersten, E. (2001) Sex-related differences in interhemispheric transmission time in the human brain. *NeuroReport* 12, 4171–4175
13 Pell, M.D. and Baum, S.R. (1997) The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. *Brain Lang.* 57, 80–99
14 Geffen, G. *et al.* (2000) Sex differences in the perception of tactile simultaneity. *Cortex* 36, 323–335
15 Kansaku, K. and Kitazawa, S. (2001) Imaging studies on sex differences in the lateralization of language. *Neurosci. Res.* 41, 333–337
16 Besson, M., and Schön, D. (2001). Comparison between language and music. In *The Biological Foundations of Music* (Zatorre, R. and Peretz, I., eds), pp. 232–259, New York Academy of Sciences
17 Buchanan, T.W. *et al.* (2000) Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cogn. Brain Res.* 9, 227–238
18 Gandour, J. *et al.* (1998) Pitch processing in the human brain is influenced by language experience. *NeuroReport* 9, 2115–2119

Mireille Besson*
Cyrille Magne
Institute for Physiological and Cognitive Neurosciences, Language and Music Group, CNRS–INPC, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France.
*e-mail: mbesson@lnf.cnrs-mrs.fr

Daniele Schön
Institut de Neurosciences Physiologiques et Cognitives and Dept of Psychology, University of Trieste, Italy.

Meeting Report

Minimizing rivalry in San Miniato

David Alais and Randolph Blake

A workshop on binocular rivalry and perceptual ambiguity was held in San Miniato, Italy, on 12–15 June, 2002.

Several dozen researchers, among them psychologists, neurophysiologists and computational modelers, came together recently to discuss perceptual ambiguity amidst delightful Tuscan surrounds. This topic refers to the dilemma facing perceptual systems when presented with optical information supporting two (or more) interpretations. Nearly always, the outcome is bistable perception, as the system alternates over time between competing interpretations. Cognitive neuroscientists have shown great interest in bistable phenomena in recent years, seeing it as a means to elucidate the neural events underlying perception and consciousness. Quite simply, as one's perceptual interpretation of an ambiguous stimulus fluctuates over time, there must be concomitant fluctuations in neural activity somewhere within the brain.

Binocular rivalry (wherein dissimilar images presented to the two eyes compete for perceptual dominance) has traditionally been the preferred paradigm for studying ambiguity. This meeting saw ambiguity explored in a variety of contexts, including: depth, structure-from-motion, figure/ground segregation, motion integration/transparency, and several others.

Intrinsic oscillators

An overarching theme concerned the degree to which all forms of perceptual ambiguity arise from a common neural mechanism. According to Jack Pettigrew (University of Queensland, Brisbane, Australia) perceptual alternations seen in many bistable stimuli might all be driven by an intrinsic oscillator. Although alternation rates vary widely among individuals, Pettigrew observed, rates for a given individual correlate with that person's cognitive style and stress levels – factors that affect their intrinsic oscillation. Consistent with a common underlying mechanism, Nava Rubin (New York University, NY, USA) showed that two seemingly unrelated forms of bistability – intermittent motion transparency in plaids and binocular rivalry – obey the same, somewhat counterintuitive principle: increasing the salience of one perceptual interpretation does not make that percept endure longer but, instead, attenuates the duration of the unchanged, weaker interpretation. At the same time, however, Hugh Wilson (York University, York, UK) cautioned against placing too much emphasis on a single underlying mechanism, pointing out that all dynamic neural systems, regardless of locus or computational purpose, will generate statistically equivalent multistable states.

Lively discussions were also triggered by Pettigrew's speculation that perceptual alternations in binocular rivalry might arise from two competing stimulus interpretations, one endorsed by each

  PII: S1364-6613(02)01987-3