

# Are there voice specific processing mechanisms? Intensity changes recruit more sensory processing resources if associated with vocal as compared to nonvocal sounds



Elizabeth Simpson<sup>1</sup>, Annett Schirmer<sup>2</sup>, and Nicolas Escoffier<sup>2</sup>

<sup>1</sup>University of Georgia, <sup>2</sup>National University of Singapore



## Abstract

Changes in the intensity of both vocal and nonvocal sounds can be emotionally relevant. However, as only vocal sounds directly reflect communicative intent, intensity change of vocal but not nonvocal sounds is socially relevant. Here we investigated whether a change in sound intensity is processed differently depending on its social relevance. To this end, participants listened passively to a sequence of vocal or nonvocal sounds that contained rare deviants which differed from standards in sound intensity. Concurrently recorded event-related potentials (ERPs) revealed a mismatch negativity (MMN) and P300 effect for intensity change. Direction of intensity change was of little importance for vocal stimulus sequences, which recruited enhanced sensory and attentional resources for both loud and soft deviants. In contrast, intensity change in nonvocal sequences recruited more sensory and attentional resources for loud as compared to soft deviants. This was reflected in markedly larger MMN/P300 amplitudes and shorter P300 latencies for the loud as compared to soft nonvocal deviants. Furthermore, while the processing pattern observed for nonvocal sounds was largely comparable between men and women, sex differences for vocal sounds suggest that women were more sensitive to their social relevance. These findings extend previous evidence of sex differences in vocal processing and add to reports of voice specific processing mechanisms by demonstrating that simple acoustic change recruits more processing resources if it is socially relevant.

## Introduction

### Importance of nonverbal communication

- Nonverbal signals (e.g. gestures, facial expressions, tone of voice) accompany language; however, both speakers and listeners are less aware of nonverbal information, compared to verbal information (Prinz, 2006; Schirmer et al., 2005)
- Despite our lack of awareness, we nonetheless possess dedicated neural mechanisms for the processing of nonverbal cues, which may call for attention if emotionally relevant (Palermo & Rhodes, 2007)
- Nonverbal signals represent important stimuli, and have primacy over the processing of other information

- Inferior temporal lobe: dedicated face processing region; perception of human faces activates this region to a larger degree than perception of animal faces or inanimate objects (Kanwisher et al., 1997; Kanwisher & Yovel, 2006)

- Superior temporal sulcus (STS): human voice processing region; responds more strongly to human vocalizations than to animal or environmental sounds (Belin et al., 2000), and vocalizations elicit even greater activation if they are emotional as compared to neutral (Ethofer et al., 2006; Grandjean et al., 2005; for a review see Schirmer & Kotz, 2006)

### Voice vs. Nonvoice Sounds

- Brain structures that mediate the primacy of vocal over nonvocal sounds have been relatively well established (for a review see Belin et al., 2004)

- However, we know little about the time course of vocal and nonvocal processing

### Sex Differences

- Females, compared to males, are more sensitive to emotional photos (e.g. snakes; Bradley et al., 2001), and nonverbal expressions (e.g. emotional faces and voices; Schirmer et al., 2005)

- Females may possess a general emotional sensitivity and/or socially relevant sensitivity

### Predictions

- Sensory processing and attention capture are facilitated for vocal as compared to nonvocal sounds
- Acoustic changes (e.g. change in volume) are more relevant to listeners when they occur in voice sounds, and therefore are processed faster and more thoroughly
- Females, compared to males, may be more sensitive to all intensity changes, and/or especially sensitive if changes are socially relevant (voice), compared to socially irrelevant (nonvoice)

## Methods

### Procedure

- Participants watched a silent movie with subtitles while passively listening to an auditory oddball sequence made up of standard sounds, comparable in their acoustic properties, and rare deviant sounds that differ from standards (either higher or lower sound intensity)
- Separate blocks were presented, one with "vocal sounds" (meaningless syllables "dada"), and another with "nonvocal sounds" (synthesized sine waves with the same temporal, fundamental frequency, and amplitude patterns as the original vocal sounds)
- We recorded two event-related potentials (ERP) from the scalp, the MMN and P300

### Mismatch negativity (MMN)

- Reflects the comparison between incoming sensory information and an existing short-lived sensory memory representation, which is updated in case of a mismatch (Näätänen et al., 2005)



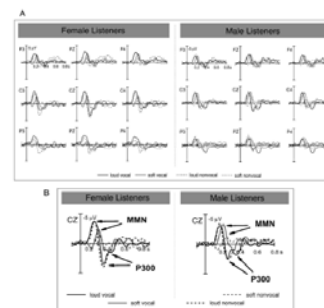
- Elicited to rare and unattended acoustic change, which may reflect a change detection mechanism (e.g., Näätänen et al., 2005; Schroeger & Wolff, 1998)
- Subtracting the ERP elicited to standards from that elicited to deviants reveals an MMN that peaks around 200 ms following stimulus onset with a fronto-central distribution (Sams et al., 1985)

### P300

- ERP that follows the MMN; considered an additional marker of change-detection
- P300 is larger in active as compared to passive paradigms, so may reflect attentional processes such as orientation towards a target (Alho et al., 1998; Friedman et al., 2001; Mecklinger et al., 1998).
- Larger P300 amplitudes associated with increased attention capture and enhanced stimulus encoding

## Results

(A) Difference waves obtained by subtracting standards from physically identical deviants. Differences waves of loud vocal sounds (solid black line), soft vocal sounds (solid gray line), loud nonvocal sounds (dotted black line) and soft nonvocal sounds (dotted gray line) are presented for female and male listeners. (B) Difference waves presented again for a single electrode site.



- Peak latencies and amplitudes were subjected to separate ANOVAs with intensity (loud/soft), stimulus (vocal/nonvocal), electrode position (anterior/posterior) and hemisphere (left/right) as repeated measures factors and sex as a between subjects factor.

### Are vocal sounds special?

- The extent to which listeners engage auditory change detection mechanisms reflected by the MMN differs as a function of social relevance
- MMN amplitudes were larger for intensity change in vocal as compared to nonvocal stimulus sequences ( $F(1,38)=11.91, p<.005$ )

### Are there differences between male and female listeners?

- MMN amplitudes were larger in female than in male listeners if a vocal stimulus sequence was interrupted by a sudden loud vocal deviant ( $F(1,38)=3.87, p<.05$ )
- Given that sex differences appeared for vocal but not for nonvocal sounds, one may infer that these sex differences reflect differential sensitivity to social information
- Females showed greater responsiveness to emotional events in general: MMN peaked earlier in females, compared to males ( $F(1,38)=5.09, p<.05$ ); P300 amplitudes were larger in females as compared to males over posterior regions ( $F(1,38)=17.91, p<.001$ )

## Discussion

These findings, consistent with previous research, suggests that human vocalizations engage dedicated processing mechanisms. We demonstrated that intensity changes recruit more sensory processing resources if they are associated with vocal as compared to nonvocal sounds. Vocal intensity deviants may call for sensory and attentional resources regardless of whether they are loud or soft, while nonvocal intensity deviants recruit comparable resources only if they are loud and hence physically salient. Intensity decreases capture attention more rapidly if they occur within vocal as compared to nonvocal sound sequences.

Vocalizations play an important role in human interactions; therefore, it is not surprising that the primacy of vocal over nonvocal signals is seen across individuals. On the other hand, there are also subtle interindividual differences in vocal processing: females appear more sensitive than males to sudden loud as compared to soft vocal expressions. It could be that while vocal sounds may be special to both men and women they are particularly special to women. Moreover, the present results support the view that sex differences in nonverbal processing reflect greater social interest and interpersonal sensitivity in females as compared to males (Hall et al., 2006; Schirmer & Kotz, 2006).

## References

- Alho, K., Winkler, I., & Escera, C. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. *Psychophysiology*, 35, 211–224.
- Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends Cogn. Sci.*, 8, 129–135.
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403, 309–312.
- Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: sex differences in picture processing. *Emotion*, 1, 300–319.
- Ethofer, T., Anders, S., Wiethoff, S., Erb, M., Herbert, C., Saur, R., Grodd, W., & Wildgruber, D. (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. *NeuroReport*, 17, 249–253.
- Friedman, D., Cyco-wicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci. Biobehav. Rev.*, 25, 355–373.
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci.*, 8, 145–146.
- Hall, J. A., Murphy, N. A., & Mast, M. S. (2006). Recall of nonverbal cues: exploring a new definition of interpersonal sensitivity. *J. Nonverbal Behav.*, 30, 141–155.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond., B Biol. Sci.*, 361, 2109–2128.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.*, 17, 4302–4311.
- Mecklinger, A., Maess, B., Opitz, B., Pfeifer, E., Cheyne, D., & Weinberg, H. (1998). A MEG analysis of the P300 in visual discrimination task. *Electroencephalogr. Clin. Neurophysiol.*, 108, 45–56.
- Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory based of afferent processes in mismatch negativity (MMN): a review of evidence. *Psychophysiology*, 42, 25–32.
- Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, 45, 75–92.
- Prinz, W. (2006). Measurement contra appearance: Oskar Pfungst examines Clever Hans. *Psychol. Rundsch.*, 57, 106–111.
- Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr. Clin. Neurophysiol.*, 62, 437–448.
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.*, 10, 24–30.
- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech: sex differences revisited. *Cogn. Brain Res.*, 24, 442–452.
- Schroeger, E., & Wolff, C. (1998). Behavioral and electrophysiological effects of task-irrelevant sound change: a new distraction paradigm. *Cogn. Brain Res.*, 7, 71–87.

This study was published in: Schirmer, A., Simpson, E., & Escoffier, N. (2007). Listen up! Processing of intensity change differs for vocal and nonvocal sounds. *Brain Research*, 1176, 103–112.

This research was supported by an Academic Research Grant (R-581-000-112/133) awarded to Annett Schirmer.