Published in final edited form as:

Brain Imaging Behav. 2009 September ; 3(3): 284–291. doi:10.1007/s11682-009-9070-7.

Mismatch Negativity to Tonal Contours Suggests Preattentive Perception of Prosodic Content

David I. Leitman^{1,2,3}, John J. Foxe², Pejman Sehatpour¹, Marina Shpaner¹, and Daniel C. Javitt^{1,2}

¹ Neuropsychiatry Program-Dept of Psychiatry, University of Pennsylvania 10th floor Gates Bldg, 3400 Spruce St. Philadelphia PA

² Department of Psychology Drexel University, Philadelphia PA

³ Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York 10962, Program in Cognitive Neuroscience, The City College of the City University of New York, 138th Street & Convent Avenue, New York, NY 10031

⁴ Departments of Psychiatry and Neuroscience, NYU School of Medicine 550 First Avenue, New York, NY

Abstract

Modulation of speech conveys information that is decoded within audio-sensory structures. For example, the termination of an utterance with a rise in pitch distinguishes statements and questions. This study evaluated the sensitivity of early auditory structures to such linguistic prosodic distinctions using mismatch negativity (MMN). MMN is a preattentive auditory event-related potential (ERP) sensitive to stimulus deviance. High-density ERP to pitch contour stimuli were collected in a passive listening oddball paradigm from 11 healthy subjects. Voltage analysis revealed significant MMN responses to declarative and interrogative oddball stimuli. Further, MMN was significantly larger to interrogative, than declarative, deviants, indicating non-symmetric brain processing. These MMNs demonstrate that pitch contour abstractions reflecting interrogative/declarative distinctions can be represented in preattentive auditory sensory memory.

Keywords

ERP; MMN; Pattern; speech; prosody

Introduction

The majority of information conveyed by speech is encapsulated within individual segments that are decoded successively into phonemes, words, and sentences. However, additional information is conveyed not only in what is said, but also in how it is stated. This suprasegmental modulation of speech contains information regarding various forms of prosodies, including, for example, whether the speaker is happy or sad (emotional prosody), or whether the utterance is a statement or a question. Although much is known about mechanisms by which segmental information in the brain is decoded into phonemes and words, relatively less is known about suprasegmental decoding mechanisms. In particular, whereas it is known that the tonal information inherent in phonemes is processed even by

Address correspondence to: David I Leitman, Ph.D., Department of Psychiatry, Brain Behavior Laboratory, University of Pennsylvania School of Medicine, 10 Gates Pavilion, 3400 Spruce ST, Philadelphia, PA. 19104, P. 215.662.7389, P. 215.662.7903, leitman@upenn.edu.

preattentive auditory mechanisms (Aaltonen, Niemi, Nyrke, & Tuhkanen, 1987), few studies (Kujala, Lepisto, Nieminen-von Wendt, Naatanen, & Naatanen, 2005; Schirmer, Striano, & Friederici, 2005) have examined preattentive processing of prosody.

These studies have focused on emotional intent or stress emphasis on phonemic pairs (Wang, Friedman, Ritter, & Bersick, 2005), and have not examined non-emotional suprasegmental prosodic distinctions like declarative versus interrogative intent.

Prior electrophysiological studies have examined preattentive processing of emotional intent using actual verbal stimuli in oddball paradigms (Kujala et al., 2005; Schirmer et al., 2005) by examining the Mismatch Negativity (MMN).

MMN is an auditory ERP elicited most commonly in the context of an auditory oddball paradigm, in which a series of standard stimuli is interrupted by an infrequent deviant stimulus. In such a paradigm, the brain automatically organizes a template reflecting invariant features of the repetitive standards. MMN then reflects the outcome of a process that compares each stimulus to the locally maintained template. MMN usually occurs with a latency of approximately 100–200 ms, which varies based upon degree of deviance. One of the characteristic features of MMN is that latency is locked to the timing of feature deviance, rather than to stimulus onset. Thus, MMN to duration deviance is delayed relative to pitch deviance, because pitch deviance can be detected very quickly after stimulus onset whereas duration deviance cannot be detected until after the offset of the deviant stimulus or the normal offset of the standard stimulus (whichever is shorter). MMN is thought to reflect the operation of neural mechanisms within auditory cortex for directing attention toward potentially significant alterations within the surrounding acoustic environment (Naatanen & Alho, 1995). MMN deficits in pitch and duration perception have been found to index aberrant audio-sensory processing in clinical populations, most notably in schizophrenia (Javitt, 2000).

As the MMN is sensitive to any deviance within the auditory environment, interpreting MMNs using actual speech has its limitations: for example, within emotion, prosodic distinctions are conveyed by changes in a constellation of highly interrelated acoustical changes such as pitch, voice intensity and voice quality, as well as temporal changes in rhythm (Juslin & Scherer, 2005). Thus, it is difficult to discern which aspect of the acoustic deviance in this complex signal is actually generating the MMN.

One key acoustical cue of speech that is used to decode prosodic intent is the contour, or trajectory of fundamental frequency (F_0), across the individual segments. In general, terminally ascending contours indicate “interrogative” intent while contours with a flat or slightly downward trajectory indicate declarative intent (Majewski & Blasdell, 1969). Suprasegmental F_0 contours alone have been shown to be sufficient for discerning interrogative or declarative intent (Majewski & Blasdell, 1969), as well as for emotional prosodic comprehension (Lakshminarayanan et al., 2003), even if underlying speech segments are masked. Further, in studies of dysprosodia resulting from RH brain lesions, Van Lancker and Sidtis (Van Lancker & Sidtis, 1992) have shown that individuals with prosodic dysfunction were also poor at using F_0 cues, suggesting that F_0 contour recognition may be among the skills necessary for proper prosodic comprehension.

There is also good reason to think that the tonal information giving rise to prosodic cues might also be processed preattentively. For example, developmental studies have suggested that prosodic cues may be detected even by prelinguistic babies, and have high interactive salience (Fernald, 1989). We hypothesized, therefore, that suprasegmental information such as declarative/interrogative distinctions would be processed automatically within low-level auditory cortical regions.

In contrast to prior MMN studies of prosody which employed real speech and examined either stress or emotional prosody, we opted for pure tonal contours that were previously shown to approximate declarative and interrogative distinctions (Pell, 1998). While real speech tokens benefit from maximum ecological validity, it is difficult to discern which aspect of the deviance generates the mismatch, as a wide variety of temporal and spectral acoustic features may differentiate the standard stimulus from the deviant. Moreover, prior research using band-passed filtered speech had suggested that F_0 contour alone is sufficient to enable one to differentiate a question from a statement (Lakshminarayanan et al., 2003). Therefore, by using tonal analogues of question statement prosody, we were able to isolate pitch as a feature that examines whether it is sufficient to generate a MMN for declarative/interrogative distinctions. More generally, this study is the first to examine whether distinctions between declarative and interrogative intent might be processed preattentively.

For the present study, the sensitivity of MMN generators to suprasegmental information was evaluated by construction of artificial stimuli that approximated the F_0 modulations signifying interrogative and declarative intent, as previously reported by Pell (Fernald, 1989; Pell, 1998). The duration and the frequency of the tones that formed each contour were matched save for the terminal tone, which pitched either upward or downward to approximate the contours of interrogative versus declarative utterances, respectively. Thus, the interrogative and declarative stimuli, each of which had a total duration of 428 ms, were identical for the first 328 ms, but deviated in frequency during the final 100 ms. Analyses of suprasegmental MMN, therefore were performed relative to onset of stimulus deviance at 328 ms.

Interrogative and declarative stimuli served as deviants in alternate runs. Given that interrogative and declarative stimuli differed in overall stimulus energy, MMN waveforms were derived by comparing ERP responses to deviant stimuli (interrogative or declarative) in one run to ERP responses to the same stimulus type in the alternate run. Thus, all MMNs in the present study were derived by subtracting like-from-like stimuli, in that the same stimulus served as both deviant and standard in alternating blocks. We hypothesized that such comparisons would elicit significant MMNs, suggesting that linguistic prosodic distinctions can be detected pre-attentively on the basis of F_0 contour alone.

MATERIALS AND METHODS

Participants

Informed consent was obtained from 14 (6 female) healthy control subjects with a mean age of 33 ± 11 yrs with no reported history of psychopathology. All subjects reported that they were right handed, had normal hearing, and were medication free at the time of testing. Three subjects (1 female) were excluded from analysis due to high levels of noise within their data. All procedures conducted were under the supervision of the local internal review board. This oddball paradigm was presented to subjects while they watched a silent movie.

Stimuli and task

Subjects were presented with two-tonal contours: an interrogative contour and a declarative contour. Each contour consisted of 5 sequential sinusoidal tones each of which had a 10ms envelope created using a Hanning window. The first four of the five tones that made up both the declarative and interrogative tonal contours were identical, matched for both duration and frequency (Table 1). Using an inter-stimulus interval (ISI) of 500 ms between tonal contours across all presentations, two types of oddball blocks were presented: an “interrogative” deviant block and a “declarative” deviant block. In the interrogative block, three declarative contours were followed by an interrogative contour in a fixed manner. In

the declarative deviant block the contour types were reversed with three interrogative standards followed by a declarative deviant. For the first four subjects, runs included an additional deviant in which the entire stimulus was shifted upward by 100 Hz (pitch deviance). However, this deviant was subsequently omitted in the interest of time. A comparison of the group average waveforms with and without these four subjects yielded no significant difference within the MMN latency window (50–150 post deviance onset) for either the declarative or interrogative subtraction waveforms (*all p's* > .58). During stimulus presentation, subjects were instructed to relax while they watched a silent movie.

The ecological validity of these stimuli was tested on a second group of ten healthy subjects [mean age: 27 ± 5 (4 male)]. All ten subjects accurately ascribed interrogative intent to the interrogative contour, and declarative intent to the declarative contour.

Four blocks of each type (declarative and interrogative) were presented each of which contained 240 standard contours and 80 deviant contours, for a total of 960 standards and 320 deviants. All comparisons were made across blocks, with the response to the deviant stimulus in one block, being compared to the response to the same stimulus in the opposite block. All tonal contours were presented binaurally at 75db (SPL) through Sennheiser HD 600 headphones. Subjects were instructed that the experiment was designed to test their passive auditory responses to tonal sequences to which they need not attend. Subjects watched a silent movie during the course of stimulus presentation.

Data Collection—High-density event-related potentials (ERP) were recorded continuously in a sound attenuated, electrically shielded booth from 128 scalp electrodes referenced to the nose with bandwidth of 0.5 to 100 Hz and digitized at a sampling rate of 500 Hz. Impedances were kept to < 5 k Ω . All recordings employed the same apparatus: Neuroscan- synamps with stimulus delivery employing STIM (Neuroscan) software.

Epochs (–200 to 700 ms relative to stimulus onset) were constructed offline. Trials with blinks and large eye movements were rejected off-line on the basis of horizontal (HEOG) and vertical (VEOG) electro-oculogram. No systematic differences in HEOG or VEOG were seen across conditions (artifact rejection window of $\pm 100 \mu\text{V}$). An artifact criterion of $\pm 100 \mu\text{V}$ was used at all other electrode sites to reject trials with excessive EMG or other noise transients from –100 ms pre-stimulus to 450 ms post-stimulus.

Accepted trials were averaged for each subject. The average number of accepted sweeps for deviant contours per condition was 250 ± 39 and was equivalent for each deviant type. For average files, baselines were corrected to zero over the –100 to 0 ms latency range. For source analysis, average files were filtered using a 0.5–45 Hz zero-phase-shift band-pass digital filter with roll-off of 24 dB/octave. This analysis was conducted using SCAN (Neuroscan) software.

Statistical Analysis

Separate statistical analyses were performed for the interrogative and declarative deviance conditions. For each deviance type, point-wise (“running”) paired t-tests (2-tailed) were calculated to detect significant differences between responses to stimuli presented as deviants versus the same stimulus (interrogative/declarative) presented as a standard. In order to protect against type I error due to multiple comparisons, we employed a significance criterion requiring at least 10 consecutive data points (= 20 ms at a 500 Hz digitization rate) to meet an 0.05 alpha criterion threshold (Guthrie & Buchwald, 1991). MMN onset and offset were defined respectively as the first and last points at which there was a statistically significant difference between standard and deviant conditions at electrode Fz, provided that this difference continued for 10 consecutive data points.

Peak amplitudes for each deviance type were determined for each subject within the overall running-t significance window. Amplitudes were defined as the most negative value for the subtraction waveform (deviant – standard) occurring within the running-t significance window. Topographical analyses of MMN distribution to interrogative versus declarative contours were conducted using ANOVA with factors of deviance type and electrode location. For lateralization analyses, lateralized amplitudes were determined by summing across 3 electrode pairs located to the left and right of Fz (approx. equivalent to locations F3 and F4). For anterior posterior analyses, 11 frontocentral electrode locations ranging from AFz to Cz were used, and analyses were corrected for non-sphericity using the Greenhouse–Geisser method. All significance levels in text are 2-tailed with preset α -level for significance of $p < 0.05$.

To estimate the sources of MMN activation, we employed a distributed linear inverse solution using Local Auto-Regressive Average (LAURA) modeling of the unknown current density in the brain (Grave de Peralta & Gonzalez, 2002). The solution space used was a realistic head model with 4024 fixed nodes that are equally distributed. LAURA solutions are capable of dealing with multiple sources of unknown location that are active simultaneously. LAURA renders solutions that best mimic the biophysical behavior of electric vector fields as it increases the amount of its sources with zero localization error while reducing the maximum error when compared to other inverse solutions. This solution was then recomputed and exported to an average brain model (Montreal Neurological Institute [MNI]) using MRICro (Rorden and Brett 2000) to be displayed as surface overlays.

RESULTS

Statistical comparisons between conditions were performed using running t-tests relative to the point at which the interrogative and declarative stimuli diverged. A comparison of the interrogative contour in the deviant versus standard position revealed a significant negative difference (MMN waveform) with an onset latency of 398 ms post-stimulus onset, corresponding to 70 ms post-deviance onset (PDO). This negative difference (MMN) had a duration of 80 ms using the running t-test method described above and a peak amplitude of $-3.5 \pm 1.2 \mu\text{V}$ ($t=3.34$, $p<0.0001$) at 112ms PDO. A comparison of the declarative contour in the deviant versus standard position also revealed a significant negative difference waveform with an onset latency of 432 ms post-stimulus onset, corresponding to 104 ms post deviance onset. The MMN waveform had a duration of 102 ms, and a peak amplitude of $-1.2 \pm 0.9 \mu\text{V}$ ($t=1.20$, $p=0.001$) at 140 ms PDO (Figure 1). Statistical comparison of onset latencies demonstrated a significantly earlier onset time for interrogative, than declarative MMN (Mean difference= 38 ± 11.0 ms, $t_{1,10}= 12.9$, $p=0.0001$).

Scalp voltage distributions of MMN are shown in Figure 2. Separate ANOVAs were conducted to compare distributions of interrogative versus declarative MMN across lateral and anterior/posterior dimensions. For lateralization analyses, MMN values were averaged for 3 pairs of electrodes centered around Fz (Table 2). Analysis was conducted using ANOVA with factors of stimulus type and hemisphere. This analysis indicated no significant hemispheric difference between deviant types ($F_{1,10} = 0.87$, $p=0.37$), but did indicate that the interrogative contour mismatch was significantly larger than its declarative counterpart ($F_{1,10} = 11.94$, $p<0.01$). Anterior/posterior analyses were performed using an ANOVA for factors of deviant type and electrode location. Electrodes that straddled the midline from posterior locations (Fz) to anterior locations (AFz) were used for this analysis. This analysis revealed a significant interaction between electrode location and stimulus type ($F_{10,100}=8.89$, $p=0.01$, $\eta^2=0.26$), reflecting a more anterior topography for MMN to the declarative versus interrogative deviance.

In general, MMN waveforms are thought to arise from auditory sensory cortex. Source localization (LAURA) analyses performed at the point of maximum amplitude for each MMN waveform revealed bilateral distributed sources that were located within superior temporal gyrus for each MMN type (Figure 3), confirming *a priori* expectations.

DISCUSSION

MMN shows well-known sensitivity to segmental aspects of speech, such as phonemic structure (Aaltonen et al., 1987). This study examined whether MMN shows similar sensitivity to suprasegmental aspects of speech, using contours constructed to approximate tonal contours associated with declarative versus interrogative utterances. A comparison of the subtraction waves across conditions found significant MMN-like activity in both conditions time-locked to the declarative/interrogative contour. These findings support the concept that, like other elements of speech, the suprasegmental contour information used to infer interrogative prosody is decoded preattentively within low-level auditory cortical regions. This result is consistent with prior studies showing sensitivity of MMN to, for example, changes in stimulus pattern (e.g. (Saarinen, Paavilainen, Schoger, Tervaniemi, & Naatanen, 1992)) as well as prior findings in emotional (Kujala et al., 2005) and stress prosody (Schirmer et al., 2005). However, this study is the first to utilize tonal contours resembling those of normal declarative and interrogative utterances, and the first to show that such information is decoded against a background of more complex spectral information.

MMN was elicited whether repetitive interrogative stimuli were presented against a background of declarative stimuli, or whether declarative stimuli were presented against a background of interrogatives. In both cases, comparisons were made between the same stimulus (i.e., interrogative/declarative) presented as a standard and the same stimulus presented as a deviant in a separate run. Thus, the spectral content of standards and deviants within each comparison was identical. Further, because of the symmetrical experimental design, the onset of deviance within each run (i.e., interrogative std/declarative deviant versus declarative std/interrogative deviant) was identical, with the contours being identical up to 328 ms and diverging thereafter.

Despite the exact symmetry in the experimental design, significant differences were observed in both the timing and amplitude of contour-elicited MMN, such that interrogative contours presented against a background of declarative contours elicited a larger MMN than declarative contours presented against a background of interrogative standards. This finding suggests that the MMN generators are responding not just to contour per se, but also to the ecological significance of the suprasegmental information contained within the contour. Commonly, for example, in both educational and social situations, individuals are able to detect when they are being asked a question, even after they have stopped paying close attention to the verbal material directed at them. The large MMN elicited by interrogative stimuli presented against a background of declarative statements may underlie the ability of questions to automatically capture attention even when the preceding declarative information has been ignored. Our findings are also consistent with the fMRI observations of Doherty et al. (Doherty, West, Dilley, Shattuck-Hufnagel, & Caplan, 2004) who found that questions with rising terminal contours elicited stronger right Heschl's gyrus (HG) and bilateral superior temporal gyrus (STG) activation than declarative utterances with descending terminal contours. The reverse contrast however of statements greater than questions revealed no regions of greater activation (Doherty et al., 2004).

Interestingly, we also observed a significant frontocentral P3 for the interrogative deviant subtraction that was not present for the declarative counterpart. This P300 had an onset

latency of 536 ms post-stimulus onset, corresponding to 208 ms post-deviance onset (PDO). This positivity had 46 ms duration using the running t-test method described above, and a peak amplitude of $1.3 \pm 1.5 \mu\text{V}$ ($t=2.69$, $p<0.007$) at 244ms PDO. In passive oddball conditions, MMN is known to trigger attentional capture, which is then indexed by later P3 potential (Naatanen & Alho, 1995). A significant P3 in the interrogative condition but not the declarative condition suggests that interrogative contours are more salient (i.e. able to capture attention) than declarative contours, in accordance with their elicitation of a larger MMN. This explanation would be consistent with the notion of the ecological utility for automatically capturing attention when hearing a question given that it naturally requires a response.

Because the present study involved only tone sequences, rather than actual sentences or abstracted speech (e.g. hummed contours), one potential criticism of the study is that the results are relevant only to processing of pitch contours in the abstract, rather than to prosody in particular. In fact, it is likely that the MMN process is not responding to the abstract concept of “interrogative” or “declarative” since MMN, in general, is thought to be cognitively impenetrable (Picton, Alain, Otten, Ritter, & Achim, 2000). However, the asymmetry of the findings do show that contours that are recognizable to the average listener as conveying either declarative versus interrogative intent do generate MMN relative to each other. This is not to say however that the MMN reflects the brain’s recognition of the “semantic” intent of these contours: rather, the MMN may signal the detection of the contour whose formation is an important milestone in a processing stream that ultimately results in such recognition. Schirmer and Kotz (Schirmer & Kotz, 2006) have proposed an object-based model analogous to models proposed for visual object recognition (Marr, 1976). Within the prosodic processing stream, acoustic features are first extracted from the speech signal and integrated into an “object” whose form is compared to a prosodic template that at latter stages of processing is linked to a particular linguistic or semantic meaning. It would make sense that such object processes may also be automatically alerting, as reflected in MMN generation. Future studies using both abstracted acoustic stimuli and more naturalistic stimuli will be needed to address these issues. Nevertheless, other studies have shown significant relationships between the ability to process language and the ability to process underlying pitch contours (Bent, Bradlow, & Wright, 2006).

Our examination of any form of hemispheric asymmetry was negative. This was somewhat unexpected given a long history of findings suggesting right hemispheric dominance for emotional prosody (Ross, 1981; Ross & Monnot, 2007; Schirmer & Kotz, 2006), as well as question/statement prosody (Blumstein & Cooper, 1974; Pell, 1998). However, a recent neuroimaging study (Doherty et al., 2004), contrasting activation for questions with rising contours versus statements with descending contours found bilateral STG activation patterns, despite greater RH laterality in frontal, and perisylvian brain regions. Our negative laterality findings could mean that hemispheric specialization does not begin at the preattentive level but it may also be that RH lateralization is not apparent in lateralization estimates based on scalp topography alone.

This study is the first to examine declarative versus interrogative intent using MMN. Despite the use of tonal contours instead of real speech, our findings are consistent with MMN to affective prosody. For example, Kujala and colleagues (Kujala et al., 2005) found that MMNs were elicited to deviant presentations of “commanding”, “sad” and “scornful” prosodic presentations of single words when contrasted with neutral standard presentations. MMN occurred with latencies ranging from 178 ms to 312 ms, which given the deviance onset between standard and deviant presentations is not inconsistent with our findings here. Thus, MMN generators may be responsive, in general, to alterations in prosody over and above sensitivity simply to the underlying spectral properties of the stimuli, suggesting that

prosodic information may be decoded, at least in part, in a preattentive, attention-independent fashion. Similar latencies of responses were also observed by Schirmer and Kotz (Schirmer & Kotz, 2003) in a MMN study of angry versus neutral prosody. Further, our response latencies using MMN and an oddball paradigm are also consistent with the findings of Wambacq et al, (Wambacq, Shea-Miller, & Abubakr, 2004) who found that non-voluntary emotional prosodic distinctions were present as early as 160 ms post-stimulus onset.

These studies used either real word sentences (Kujala et al., 2005) or pseudo single words (Schirmer et al., 2005). In both cases, the MMN differences present could be do to complex differences in temporal aspects of prosody such as jitter, pause proportion or attack, voice quality differences such as spectral energy, voice intensity differences, and pitch differences. Here, by using tonal contours, which approximate F0 changes in declarative versus interrogative distinctions, we revealed similar MMN differences to these prior studies that can only be attributable to frequency differences. This finding suggests that pitch change alone may be sufficient for preattentive prosodic processing.

Gottselig et al.,(Gottselig, Brandeis, Hofer-Tinguely, Borbely, & Achermann, 2004) recently demonstrated that practicing discrimination of tone sequences can increase MMN amplitude to the practiced sequence relative to an unpracticed sequence. Similar MMN enhancements in response to training can be seen with phonemes(Kraus et al., 1995) as well as with short musical sequences(Lappe, Herholz, Trainor, & Pantev, 2008). In the present study, however, discrimination was not practiced during the study. Instead, the asymmetric response pattern to interrogative versus declarative sequences must take advantage of either innate or learned discriminations that develop for most people over the course of development.

Deficits in MMN generation to simple frequency and duration deviances have been reported in pathological conditions, such as schizophrenia, and have been shown to correlate with impaired discrimination of basic tonal deviances (Javitt, 2000). Furthermore, deficits in MMN generation contribute significantly to poor psychosocial outcome in both patients with schizophrenia (Light, Swerdlow, & Braff, 2007) and normal volunteers(Light & Braff, 2005).

In addition, patients with schizophrenia show impaired decoding of auditory prosodic information (D. I. Leitman et al., 2005) linked to neurostructural abnormalities in primary auditory regions (D. I. Leitman et al., 2007), although neural correlates of prosodic impairment still require more investigation. The sensitivity of MMN to tonal contours approximating prosodic distinctions suggests that MMN can be used as well to probe neural bases of prosodic function in both normative and pathological populations. Although MMN to prosodic contours has not yet been evaluated in schizophrenia, present results would predict impaired MMN to prosodic contours related to impaired prosodic processing in schizophrenia. The larger MMN to interrogative-like stimuli embedded among more frequent declarative contours than to declarative-among-interrogative contours suggests fundamental differences in contour salience even at early stages of auditory information processing.

In conclusion, the ability to detect tonal changes in speech is a key aspect of social interaction. The present findings suggest that MMN may be useful for assessing the locus of dysfunction in conditions such as schizophrenia and autism that are associated with impaired detection of prosodic information, particularly if this prosodic dysfunction is linked to impaired auditory perception.

Acknowledgments

This work was supported in part by NIMH grants NRSA F1-MH067339 (DIL), K02 MH01439 and R01 MH49334 (DCJ), R37 MH49334 and NS30029 (JJF), and a Translational Research Scientist Award from the Burroughs Wellcome Fund (DCJ).

References

- Aaltonen O, Niemi P, Nyrke T, Tuhkanen M. Event-related brain potentials and the perception of a phonetic continuum. *Biol Psychol.* 1987; 24(3):197–207. [PubMed: 3663795]
- Bent T, Bradlow AR, Wright BA. The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *J Exp Psychol Hum Percept Perform.* 2006; 32(1):97–103. [PubMed: 16478329]
- Blumstein S, Cooper WE. Hemispheric processing of intonation contours. *Cortex.* 1974; 10(2):146–158. [PubMed: 4844467]
- Cherry, C. On human communication: a review, a survey and a criticism. 3. Cambridge, Mass: MIT Press; 1978.
- Doherty CP, West WC, Dilley LC, Shattuck-Hufnagel S, Caplan D. Question/statement judgments: an fMRI study of intonation processing. *Hum Brain Mapp.* 2004; 23(2):85–98. [PubMed: 15340931]
- Fernald A. Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development.* 1989; 60(6):1497–1510. [PubMed: 2612255]
- Gottselig JM, Brandeis D, Hofer-Tinguely G, Borbely AA, Achermann P. Human central auditory plasticity associated with tone sequence learning. *Learn Mem.* 2004; 11(2):162–171. [PubMed: 15054131]
- Grave de Peralta R, Gonzalez A. Comparison of algorithms for the localization of focal sources: evaluation with simulated data and analysis of experimental data. *Int J Bioelectromagn.* 2002; 4(1) online journal.
- Guthrie D, Buchwald JS. Significance testing of difference potentials. *Psychophysiology.* 1991; 28(2): 240–244. [PubMed: 1946890]
- Javitt DC. Intracortical mechanisms of mismatch negativity dysfunction in schizophrenia. *Audiol Neurotol.* 2000; 5(3–4):207–215. [PubMed: 10859415]
- Juslin, P.; Scherer, K. Vocal expression of affect. In: Harrigan, J.; Rosenthal, R.; Scherer, K., editors. *The new Handbook of methods in nonverbal behavior research.* New York: Oxford University Press; 2005. p. 65–135.
- Kraus N, Mcgee T, Carrell TD, King C, Tremblay K, Nicol T. Central Auditory-System Plasticity Associated with Speech-Discrimination Training. *Journal of Cognitive Neuroscience.* 1995; 7(1): 25–32.
- Kujala T, Lepisto T, Nieminen-von Wendt T, Naatanen P, Naatanen R. Neurophysiological evidence for cortical discrimination impairment of prosody in Asperger syndrome. *Neurosci Lett.* 2005; 383(3):260–265. [PubMed: 15885908]
- Lakshminarayanan K, Ben Shalom D, van Wassenhove V, Orbelo D, Houde J, Poeppel D. The effect of spectral manipulations on the identification of affective and linguistic prosody. *Brain Lang.* 2003; 84(2):250–263. [PubMed: 12590914]
- Lappe C, Herholz SC, Trainor LJ, Pantev C. Cortical plasticity induced by short-term unimodal and multimodal musical training. *Journal of Neuroscience.* 2008; 28(39):9632–9639. [PubMed: 18815249]
- Leitman D, Hoptman M, Foxe JJ, Wylie GR, Nierenberg J, Jalbkowski M, et al. The Neural Substrates of Impaired Prosodic Detection in Schizophrenia and its Sensorial Antecedents. *Am J Psychiatry.* 2007; 164(3):1–9. [PubMed: 17202533]
- Leitman DI, Foxe JJ, Butler PD, Saperstein A, Revheim N, Javitt DC. Sensory contributions to impaired prosodic processing in schizophrenia. *Biol Psychiatry.* 2005; 58(1):56–61. [PubMed: 15992523]
- Light GA, Braff DL. Mismatch negativity deficits are associated with poor functioning in schizophrenia patients. *Arch Gen Psychiatry.* 2005; 62(2):127–136. [PubMed: 15699289]

- Light GA, Swerdlow NR, Braff DL. Preattentive sensory processing as indexed by the MMN and P3a brain responses is associated with cognitive and psychosocial functioning in healthy adults. *J Cogn Neurosci*. 2007; 19(10):1624–1632. [PubMed: 18271737]
- Majewski W, Blasdel R. Influence of fundamental frequency cues on the perception of some synthetic intonation contours. *J Acoust Soc Am*. 1969; 45(2):450–457. [PubMed: 5779530]
- Marr D. Representation and recognition and organization of three dimensional shapes. *Philosophical Transactions of the Royal Society of London*. 1976; 275:483–524. [PubMed: 12519]
- Naatanen R, Alho K. Mismatch negativity--a unique measure of sensory processing in audition. *Int J Neurosci*. 1995; 80(1–4):317–337. [PubMed: 7775056]
- Pell MD. Recognition of prosody following unilateral brain lesion: influence of functional and structural attributes of prosodic contours. *Neuropsychologia*. 1998; 36(8):701–715. [PubMed: 9751436]
- Picton TW, Alain C, Otten L, Ritter W, Achim A. Mismatch negativity: different water in the same river. *Audiol Neurotol*. 2000; 5(3–4):111–139. [PubMed: 10859408]
- Ross ED. The aprosodias. Functional-anatomic organization of the affective components of language in the right hemisphere. *Arch Neurol*. 1981; 38(9):561–569. [PubMed: 7271534]
- Ross ED, Monnot M. Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain Lang*. 2007
- Saarienen J, Paavilainen P, Schoger E, Tervaniemi M, Naatanen R. Representation of abstract attributes of auditory stimuli in the human brain. *Neuroreport*. 1992; 3(12):1149–1151. [PubMed: 1493229]
- Schirmer A, Kotz SA. ERP evidence for a sex-specific Stroop effect in emotional speech. *J Cogn Neurosci*. 2003; 15(8):1135–1148. [PubMed: 14709232]
- Schirmer A, Kotz SA. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn Sci*. 2006; 10(1):24–30. [PubMed: 16321562]
- Schirmer A, Striano T, Friederici AD. Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport*. 2005; 16(6):635–639. [PubMed: 15812323]
- Van Lancker D, Sidtis JJ. The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not created equal. *J Speech Hear Res*. 1992; 35(5): 963–970. [PubMed: 1447930]
- Wambacq II, Shea-Miller KJ, Abubakr A. Non-voluntary and voluntary processing of emotional prosody: an event-related potentials study. *Neuroreport*. 2004; 15(3):555–559. [PubMed: 15094522]
- Wang J, Friedman D, Ritter W, Bersick M. ERP correlates of involuntary attention capture by prosodic salience in speech. *Psychophysiology*. 2005; 42(1):43–55. [PubMed: 15720580]

Figure 1. Scalp grand average waveforms

The first two columns illustrate the waveforms elicited by the interrogative (left) and declarative (right) contours when presented either as standard (blue trace) or deviant (red trace) stimuli within the MMN stimulation paradigm at indicated electrodes. The final column illustrates the deviant minus standard subtraction waveforms for the interrogative (black) and declarative (grey) stimulus types. The horizontal yellow bars lying on the x-axis in the first two columns reflect the period of significant difference between the standard and deviant presentations of the contours. The black and grey arrows under the subtraction waveforms reflect the maximum peak of the interrogative and declarative contour subtractions respectively.

Voltage Topography

Figure 2. Topography of MMN activity

Voltage maps illustrating distribution of activity for interrogative (top) and declarative (bottom) difference waveforms at latency of peak MMN stimulus onset.

Figure 3. Source analysis of MMN activity

Distributed inverse solution for MMN generators showing bilateral sources within superior temporal plane (primary and secondary auditory cortex). Solutions were determined by Local Auto-Regressive Average (LAURA) (Grave de Peralta & Gonzalez, 2002). modeling. A graph of the global field power for each of the deviant conditions is presented to the left the source analysis figures.

Table 1

Contour characteristics of stimuli used in this study (adapted from Pell [9])

Contour	Hz	msec	Hz	msec	Hz	msec	Hz	msec	Total Duration
Interrogative	227	104	211	107	183	117	320	100	428
Declarative	227	104	211	107	183	117	162	100	428

Table 2

Peak MMN amplitude by hemisphere (n=11)

Variable	Peak Amplitude(μ V) – mean (sd)	
	Left hemisphere	Right hemisphere
Declarative contour (LH)	1.8 (.7)	1.6 (.9)
Interrogative contour (LH)	3.5 (1.8)	3.3 (1.1)