Available online at www.sciencedirect.com

SCIENCE DIRECT®

Brain and Language 85 (2003) 396–401

Brain and Language

www.elsevier.com/locate/b&l

The brain generates its own sentence melody: A Gestalt phenomenon in speech perception

Christoph S. Herrmann,[a,b,*] Angela D. Friederici,[a] Ulrich Oertel,[a] Burkhard Maess,[a] Anja Hahne,[a] and Kai Alter[a]

[a] *Max Planck Institute of Cognitive Neuroscience, P.O. Box 500 355, 04303 Leipzig, Germany*
[b] *Otto-von-Guericke University Magdeburg, Magdeburg, Germany*

Accepted 21 February 2003

Abstract

Brain processes underlying spoken language comprehension comprise auditory encoding, prosodic analysis and linguistic evaluation. Auditory encoding usually activates both hemispheres while language-specific stages are lateralized: analysis of prosodic cues are right-lateralized while linguistic evaluation is left-lateralized. Here, we investigated to what extent the absence of prosodic information influences lateralization. MEG brain-responses indicated that syntactic violations lead to early bi-lateral brain responses for syntax violations. When the pitch of sentences was flattened to diminish prosodic cues, the brain's syntax response was lateralized to the right hemisphere, indicating that the missing pitch was generated automatically by the brain when it was absent. This represents a Gestalt phenomenon, since we perceive more than is actually presented.

Keywords: Prosody; MEG; Syntax; ELAN; Pitch; Lateralization; Brain; Gestalt

1. Introduction

Auditory encoding, prosodic analysis, and linguistic evaluation constitute the brain processes required to understand spoken language. Prosodic cues, like sentence melody or intonation, are probably evaluated by subsequent processes of linguistic analysis, such as syntactic parsing and semantic integration. Both of these types of analyses result in individual event-related potentials (ERPs) in the human electroencephalogram (EEG): syntactic violations lead to the so-called early left anterior negativity (ELAN) (Friederici, 1997). However, the early syntax-related anterior negativity, although being lateralized to the left in a number of studies (Friederici, Mecklinger, & Hahne, 1996; Friederici, Pfeifer, & Hahne, 1993; Hahne & Friederici, 1999), not always demonstrates a clear left maximum but sometimes shows a bi-lateral distribution (Friederici, von Cramon, & Kotz, 1999; Knösche, Maess, & Friederici, 1999). An early anterior negativity with a right hemisphere dominance (ERAN) was found for violations of syntax-like disharmonic patterns in music (Kölsch, Gunter, & Friederici, 2000). Musical features such as frequency, rhythm, and intonation also appear in natural speech and are called the prosodic cues of spoken language which have been shown to be important factors of speech comprehension (Kimberly, Lindfield, Wingfield, & Goodglass, 1999) at the segmental and suprasegmental level. Pitch (F0 frequency), especially, influences the correct classification of spoken words in languages such as German and English (Pell & Baum, 1997). These processes of prosodic analysis are believed to be mediated by right-hemispheric mechanisms (Pell, 1999). The right prefrontal cortex was shown to support pitch discrimination in speech syllables (Zatorre, Evans, Meyer, & Gjedde, 1992). At the suprasegmental level pitch modulations appear to affect brain activation in the right more than in the left hemisphere (Lattner, Maess, Wang, Friederici, & Alter, 2001). Thus, it is conceivable that the degree to which

---

[*] Corresponding author. Fax: +49-341-9940-204.
E-mail address: herrmann@cns.mpg.de
URL: http://www.uni-magdeburg.de/cherrman (C.S. Herrmann).

0093-934X/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0093-934X(03)00054-3

the early syntax-related anterior negativity is lateralized depends upon the parsing system: if the latter mainly relies upon syntactic processes it would be lateralized to the left hemisphere, if, however, it additionally considers prosodic information it would be lateralized to the right hemisphere.

The interdependence of these two types of information was demonstrated in a recent electrophysiological study indicating that prosodic information is used to guide early syntactic structure building (Steinhauer, Alter, & Friederici, 1999).

In previous MEG experiments, we found that syntactic violations in spoken language lead to an early anterior syntax component, distributed bi-laterally in some experiments (Herrmann, Oertel, Maess, Wang, & Friederici, 2000; Knösche et al., 1999) but lateralized to the left hemisphere in others (Friederici, Wang, Maess, Herrmann, & Oertel, 2000). A recent dipole analysis of the magnetic data for the early syntax-related component (ELAN) revealed a fronto-lateral and a temporal dipole in each hemisphere with a tendency to larger amplitudes within the left hemisphere (Friederici et al., 2000). In that study, dipoles were constrained by anatomical locations obtained in an earlier fMRI experiment (Meyer, Friederici, & von Cramon, 2000). From these studies we can conclude that both the left and right hemisphere are involved in auditory sentence processing. This conclusion is supported by recent fMRI studies investigating spoken sentence comprehension (Friederici, Meyer, & von Cramon, 2000; Meyer et al., 2000; Müller et al., 1997). The specific contribution of the right hemisphere, however, is still to be determined.

The present experiment was designed to test the hypothesis that the degree of lateralization of the early syntax-related negativity depends on the reliance on prosodic cues available in the auditory sentence input.

For the experiment reported here, we flattened the pitch of the sentences used in a previous experiment (Herrmann et al., 2000). With this procedure, optimal prosodic cues such as the normally present global pitch contour, i.e., the typical rising and falling F0 pattern over the whole sentence, were eliminated resulting in monotonously sounding sentence material. Subjects had to listen to correct and syntactically incorrect sentences and judge their grammatical correctness.

According to the assumption that the processing of prosodic cues is carried out by the right hemisphere (Pell, 1999), we expected our pitch-flattening to decrease the right-hemispheric component of the magnetic syntax-related component.

2. Methods

Eleven student subjects (4 female), aged 19–29 (mean age 23.2), were investigated. All subjects were right-handed (laterality index 100). All subjects gave written informed consent and showed no signs of neurological, psychiatric, or hearing disorders. Two subjects of an initial set of 13 subjects had to be excluded from analysis due to artifacts. Three types of experimental sentences were presented. Correct sentences comprised a noun phrase, an auxiliary and a past participle (e.g., ‘Der Fisch wurde geangelt’ / ‘The fish was caught’). In order to avoid subjects from judging a violation (see below) depending on the preposition preceding the final word, a second class of correct sentences was presented as fillers which were not analyzed. These filler sentences were of the form ‘Der Fisch wurde im See geangelt’ (free translation: ‘The fish was caught in the lake’). Note that German is a verb final language resulting in subject-object-verb word order leading to a literal translation like ‘The fish was in the lake caught.’

Syntactically incorrect sentences were presented containing a phrase structure violation. In these sentences a preposition (‘in the’) appeared after the auxiliary (‘was’) and was directly followed by a past participle: ‘Der Fisch wurde im geangelt’ (literal translation: ‘The fish was in the caught’, violation underlined, free translation: ‘The fish was caught in the’). Since the preposition obligatorily requires a subsequent noun phrase (i.e., lake) to make up the prepositional phrase the above sentence represents a phrase category violation because the parser receives a verb instead of a noun.

ERFs were calculated for the critical (final) word. The sentences were identical to the previous MEG study (Herrmann et al., 2000). The stimuli underwent an automatically performed analysis by means of the PRAAT speech editor (Boersma & Weenink, 2000). Intensity, duration, and spectral properties were analyzed and maintained. The pitch contour was extracted (using autocorrelation) and subsequently flattened: i.e., all F0 values were set to 180 Hz, which was the average value of the female voice and a PSOLA (pitch synchronous overlap and add) resynthesis of the whole speech signal was performed so that the new flat F0 contour was combined with the previous signal parameters. Subjects' hearing thresholds were determined and sentences were presented 50 dB above. MEG was recorded with a BTI 148 channel whole-head system (MAGNES WHS 2500). Horizontal and vertical EOG was registered with four additional EEG electrodes. Data were sampled at 508.63 Hz (on-line 0.1 Hz analog high-pass and 100 Hz low-pass filtering) and digitally off-line filtered with a 2 Hz high-pass and a 10 HZ low-pass filter to avoid baseline correction. The subjects' head positions were recorded via 5 coils and headshapes were digitized with a 3D digitizer.

Averaging epochs lasted from 100 ms before to 500 ms after stimulus onset. All epochs were at first automatically and then manually inspected for artefacts and rejected if eye-movement artefacts or sensor drifts were detected. For automatic detection, we computed the standard deviation in a moving time window and epochs were rejected if a threshold was exceeded. EOG electrodes and MEG channels were checked with thresholds of 30 μV and 1100 fT with window sizes of 200 ms and 3 s, respectively. Also, if the min–max value of any sensor exceeded a threshold of 3000 fT it was rejected. In case adjacent sensors (distance 40 mm) showed mean absolute correlations of the magnetic field strengths of less than 0.75 they were rejected as artefactual. Two subjects had to be excluded from further analysis due to artefacts. Individual subjects' data were transformed to a standard gradiometer before further analysis (averaging across blocks, sessions, and subjects as well as computing the statistics) to avoid distortions due to different head sizes. The surface derivative of the event-related fields (ERFs) was computed to obtain one maximum over the location of the source, instead of a dipolar field distribution.

Repeated-measures ANOVAs with factors hemisphere (left, right) and condition (correct, incorrect) were conducted to assess the effects of the experimental variables on the dependent variable for the ELAN time window (120–200 ms). Statistics were Huynh–Feldt corrected.

3. Results

Fig. 1 shows one of the sentences in the time domain (top) and its F0 frequency over time before (middle) and after (bottom) the pitch had been flattened.

An ANOVA of the surface derivative of the event-related magnetic fields in the time interval 120–200 ms after the critical word yielded a significant main effect of condition ($F(1,10) = 10.98, p < .01$), indicating larger amplitudes for syntactic violations than for correct sentences (cf. Fig. 2 top). In addition, the interaction condition × hemisphere was significant ($F(1,10) = 7.03$, $p < .05$). Post-hoc comparisons in each hemisphere revealed that the factor condition yielded a significant effect only over the right hemisphere ($F(1,10) =$

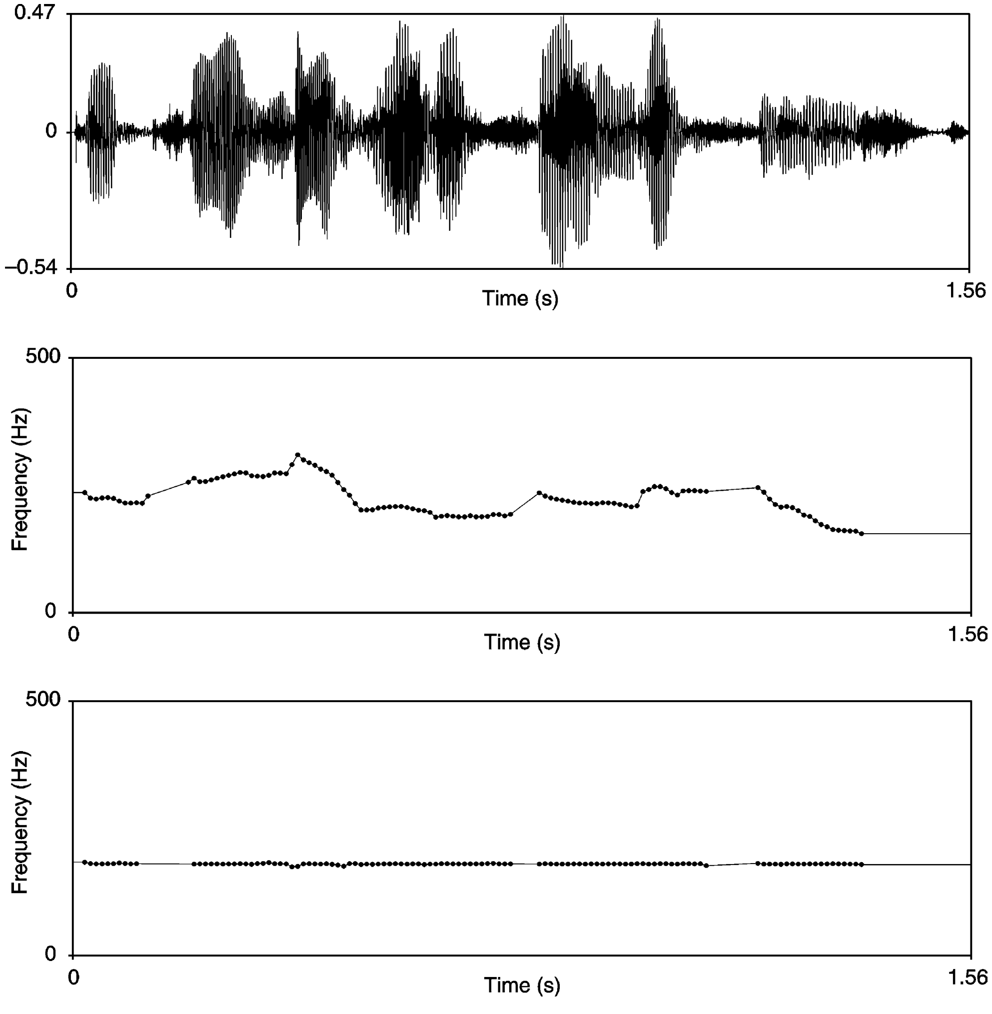

Fig. 1. Audio signal of sentence material (top) and frequency of F0 before (middle) and after pitch flattening (bottom).

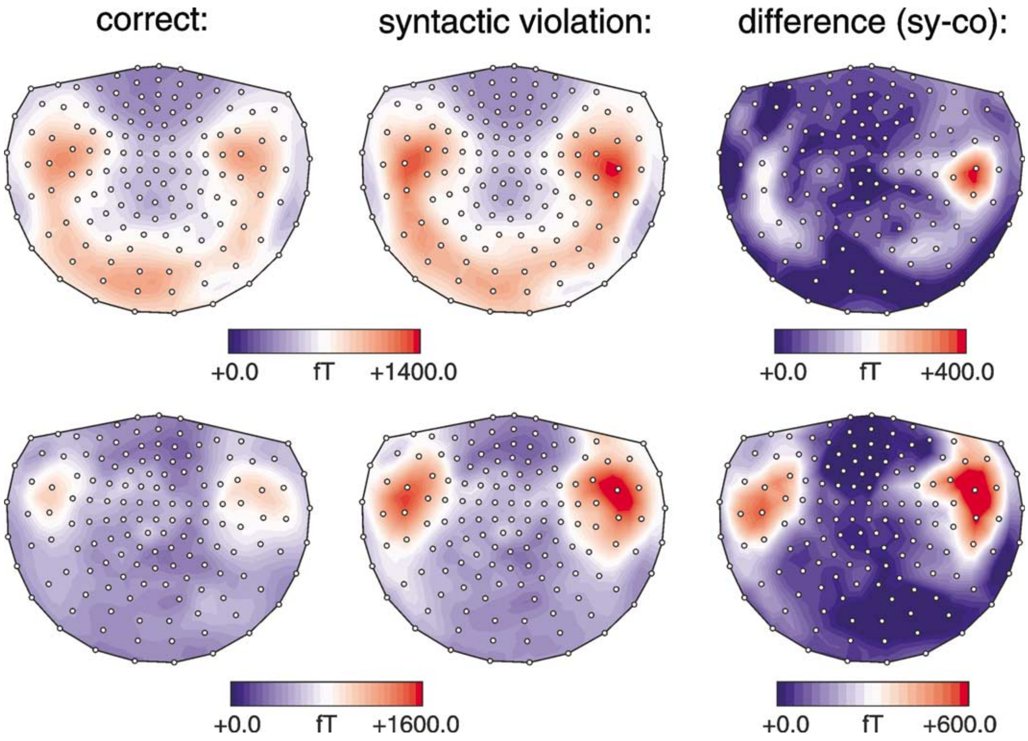

Fig. 2. Topographical maps of the surface derivative of correct and syntax condition as well as the difference (syntax—correct) for pitch-flattened sentences (top row) and regular sentences (bottom row). The brain's response to syntactical violations of spoken language (120–200 ms) is lateralized to the right hemisphere for pitch-flattened sentences, as seen in the difference maps (top right). Sentences with regular pitch information yield a symmetrical syntax response (bottom right). The magnetic field density is measured in femto Tesla per meter (fT/m). The overall amplitudes of all responses to pitch-flattened sentences (top) decreased as compared to sentences with normal prosodic cues (bottom).

$13.73, p < .005$, 623 fT/m (correct) vs. 716 fT/m (syntax)) but not over the left hemisphere (633 fT/m (correct) vs. 667 fT/m (syntax)). The stronger activation of the right hemisphere for flat-pitch sentences can bee seen in Fig. 2 (top row).

An ANOVA on the N100 at sentence onset yielded no significant differences between hemispheres.

To compare the previous MEG experiment (Herrmann et al., 2000) without pitch flattening with the present data, we re-analyzed the previous data. We reduced the original number of subjects ($n = 16$) to match the present investigation ($n = 11$) keeping those subjects which participated in both experiments ($n = 5$). We then computed the surface derivatives and reran ANOVAs and post-hoc analyses.

This reanalysis of the previous experiment yielded a significant main effect of condition ($F(1, 10) = 68.52$, $p < .0001$) but no significant interaction of condition x hemisphere ($F(1, 10) = 0.00, p = .95$). Post-hoc comparisons yielded significant main effects of condition over left ($F(1, 10) = 27.50, p < .0005$, 472 fT/m (correct) vs. 640 fT/m (syntax)) and right hemisphere ($F(1, 10) = 35.49, p < .0001$, 492 fT/m (correct) vs. 657 fT/m (syntax)), indicating symmetric processing of syntactic violations when prosodic cues are present. This symmetric activation of both hemispheres is illustrated in Fig. 2 (bottom).

4. Discussion

In line with our hypothesis, pitch flattening affected the laterality of the magnetic early syntax-related component, however, in the opposite direction. In the present study we found a condition effect only over the right hemisphere, while condition effects were found in both hemispheres for sentences with normal prosody (Herrmann et al., 2000).

Following the hypothesis that the right hemisphere supports prosodic processes, the direction of the change in laterality seems to be counterintuitive. A stronger right-hemispheric syntax-related magnetic component was found in the absence of sufficiently well-formed prosodic cues. Pitch flattening leads to less optimized prosodic characteristica for the sentence material. Thus, the increased right-hemispheric involvement for sentences without the full bundle of appropriate prosodic properties may reflect additional processes necessary to deal with prosodically non-optimal language input. Once the processing system enters into a syntactically driven phrase category clash, the detection of such a syntactic mismatch results in an early left hemispheric or bi-lateral magnetic component (Friederici et al., 2000). If this kind of auditory sentence processing and syntactic mismatch detection is in parallel aggrevated by missing prosodic cues the processing system seems to activate additional resources. This supplementary neural mobilization in our study has been detected by an increase of activity in the right hemisphere. Note, however, that the overall amplitudes of all responses to pitch-flattened sentences decreased as compared to sentences with normal prosodic cues leading also to decreased differences between conditions. The effect of prosodic variation on the early syntactic negativity strongly suggests an interaction between syntactic and prosodic information during initial processes. The notion that prosody interacts with syntax has been recently supported by theoretical considerations (Fodor, 1998) as well as by empirical data on auditory sentence processing (Steinhauer et al., 1999). In the latter study it was demonstrated that prosodic cues are immediately used for the processing of the syntactic structure. When processing a syntactically ambiguous sentence the system uses prosodic information for disambiguation and initial phrase structure building. This finding as well as our result indicate an early interaction of prosodic and syntactic information during auditory sentence processing. Interestingly, the assumption that the processing of pitch flattened sentence stimuli per se has a *general* impact on the right-hemispheric activation can be ruled out by the analysis of the behavior of early auditory components such as the N100 at the sentence onset: The N100 at this position does not show significant differences between hemispheres. The shift of activation to the right hemisphere takes place only in the case of insufficient prosodic information accompanied by a syntactic mismatch. Thus, it appears that the increase in right hemispheric activation is located at a linguistic level. It seems plausible to assume that the increase of right-hemispheric activation reflects the generation of the missing pitch to support linguistic analysis.

Our findings are in line with ideas from Gestalt Psychology in that the whole is more than the sum of its parts. In our case the missing prosody of a spoken sentence was filled in by the human brain to ease syntactic interpretation. Similar phenomena are well known from the visual domain where humans perceive squares when only fragments of the square are presented, as in case of the Kanizsa square (Herrmann & Bosch, 2001). In that case neurons, which usually detect lines in the visual field, show activity without an actual line being presented in the neuron's receptive field (Grosof, Shapley, & Hawken, 1993). A similar phenomenon seems to be at work when the missing prosody is filled in.

Further research must show to what extent the observed right hemisphere activation is due to processes which add the missing prosodic information to the incoming signal or to prosodic information processing at the suprasegmental level in general.

References

Boersma, P., & Weenink, D., 2000. Praat: Doing phonetics by computer. PRAAT. Institute of Phonetic Sciences, University of Amsterdam, Available http://www.fon.let.uva.nl/praat/.

Fodor, J. (1998). Leanring to parse? *Journal of Psycholinguistic Research, 27*, 285–319.

Friederici, A. D. (1997). Neurophysiological aspects of language processing. *Clinical Neuroscience, 4*, 64–72.

Friederici, A. D., von Cramon, D. Y., & Kotz, S. A. (1999). Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. *Brain, 122*, 1033–1047.

Friederici, A. D., Mecklinger, A., & Hahne, A. (1996). The temporal structure of syntactic parsing: Early and late event-related brain potential effects elicited by syntactic anomalies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1219–1248.

Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language, 74*, 289–300.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1*, 183–192.

Friederici, A. D., Wang, Y., Maess, B., Herrmann, C. S., & Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. *Human Brain Mapping, 11*, 1–1.

Grosof, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque V1 neurons can signal 'illusory' contours. *Nature, 365*, 550–552.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*, 194–205.

Herrmann, C.S., Oertel, U., Wang, Y., Maess, B., & Friederici, A. D (2000). Noise affects auditory and linguistic processing differently: An MEG study. *Neuroreport, 11*, 227–229.

Herrmann, C. S., & Bosch, V. (2001). Gestalt perception modulates early visual processing. *Neuroreport, 12*(5), 901–904.

Kimberly, L., Lindfield, K., Wingfield, A., & Goodglass, H. (1999). The role of prosody in the mental lexicon. *Brain and Language, 68*, 312–317.

Knösche, T. B., Maess, B., & Friederici, A. D. (1999). Processing of syntactic information monitored by brain surface current density mapping based on MEG. *Brain Topography, 12*(2), 75–87.

Kölsch, S., Gunter, T. C., & Friederici, A. D., et al. (2000). Brain indices of music processing: Non-musician are musical. *Journal of Cognitive Neuroscience, 12*, 520–541.

Lattner, S., Maess, B., Wang, Y., Friederici, A. D., & Alter, K., (2001). Human auditory processing of pitch-altered speech. In: van Dommelen, W.A., Fretheim, T. (Eds.), *Nordic Prosody. Proceedings of the 8th Conference, Trondheim 2000* (pp. 165–171). Frankfurt am Main: Peter Lang.

Meyer, M., Friederici, A. D., & von Cramon, D. Y. (2000). Neurocognition of auditory sentence comprehension: Event-related fMRI reveals sensitivity to syntactic violations and task demands. *Cognitive Brain Research, 9*, 19–33.

Müller, R. A., Rothermel, R. D., Behen, M. E., Muzik, O., Mangner, T. J., & Chugani, H. T. (1997). Receptive and expressive language activations for sentences: A PET study. *NeuroReport, 8*, 3767–3770.

Pell, M. (1999). Fundamental frequency encoding of linguistic and emotional prosody by hemisphere-damaged speakers. *Brain and Language, 69*, 161–192.

Pell, M., & Baum, S. (1997). Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody. *Brain and Language, 57*, 195–214.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2*, 191–196.

Zatorre, R.J., Evans, A.C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256*, 846–849.