Specific Neural Traces for Intonational Discourse Categories as Revealed by Human-evoked Potentials

Joan Borràs-Comes¹, Jordi Costa-Faidella², Pilar Prieto¹,³, and Carles Escera²

Abstract

The neural representation of segmental and tonal phonological distinctions has been shown by means of the MMN ERP, yet this is not the case for intonational discourse contrasts. In Catalan, a rising–falling intonational sequence can be perceived as a statement or as a counterexpectational question, depending exclusively on the size of the pitch range interval of the rising movement. We tested here, using the MMN, whether such categorical distinctions elicited distinct neurophysiological patterns of activity, supporting their specific neural representation. From a behavioral identification experiment, we set the boundary between the two categories and defined four stimuli across the continuum. Although the physical distance between each pair of stimuli was kept constant, the central pair represented an across-category contrast, whereas the other pairs represented within-category contrasts. These four auditory stimuli were contrasted by pairs in three different oddball blocks. The mean amplitude of the MMN was larger for the across-category contrast, suggesting that intonational contrasts in the target language can be encoded automatically in the auditory cortex. These results are in line with recent findings in other fields of linguistics, showing that, when a boundary between categories is crossed, the MMN response is not just larger but rather includes a separate subcomponent.

INTRODUCTION

A series of studies have indicated that segmental and tonal phonological distinctions can be represented in preattentive auditory sensory memory. However, there is no conclusive evidence with respect to the neurophysiological representation of intonational discourse contrasts (i.e., between statements and questions), and no previous research has dealt with the processing of intonational within-category and across-category contrasts. In this article, we report a study that uses the auditory MMN ERP to test the native perception of within-category and across-category intonational contrasts between statement and question interpretations in Catalan. We hypothesize that discrete intonational information as well as discrete phonological information can be represented through symbolic memory traces (in contrast to mere acoustic memory traces) in the brain.

The MMN component is a negative deflection of the auditory ERP, occurring between 100 and 250 msec after the onset of a stimulus violating an established acoustic regularity. Traditionally, it is obtained by subtracting the ERP to a standard (STD) stimulus from that to a deviant (DEV) stimulus that is presented in the same block of trials. The MMN is generally elicited in nonattentive conditions and typically argued to reflect preattentive detection of auditory changes and higher-level cognitive processes in the auditory system (Pulvermüller & Shtyrov, 2006; Näätänen, 2001). Following Näätänen (2001), the MMN reflects the early access to stored linguistic representations and indicates the match or mismatch between a stimulus and its corresponding symbolic memory trace in the brain. According to Pulvermüller and Shtyrov (2006), the MMN for language stimuli is composed of at least two parts: a part that reflects the automatic detection of a sound change and a part that reflects the activation of cortical cell assemblies forming the long-term memory traces for learned cognitive representations (see Fournier, Gussenhoven, Jensen, & Hagoort, 2010, for a review of the studies on the lateralization of tonal and intonational pitch processing).

The MMN has been successfully applied in studies of segmental phonetic and phonological analysis (e.g., Sharma & Dorman, 2000; Winkler et al., 1999; Dehaene-Lambertz, 1997; Näätänen et al., 1997) and abstract phonological features (Eulitz & Lahiri, 2004; Phillips et al., 2000; for a review, see Näätänen, Paavilainen, Rinne, & Alho, 2007; Näätänen, 2001). Näätänen et al. (1997) suggested that the identification of the DEV as a native-language vowel enhanced the MMN amplitude, that is, the phonological representation of a vowel sound can be probed with the mismatch process. Native across-category consonant contrasts also elicit a significant MMN compared with nonnative contrasts.
or within-category contrasts (Dehaene-Lambertz, 1997). A series of studies have demonstrated that acoustic contrasts that cross a phonemic boundary lead to larger MMN responses than comparable acoustic contrasts that do not (Phillips et al., 2000; Sharma & Dorman, 1999; Aaltonen, Eerola, Hellström, Uusipaikka, & Lang, 1997; Dehaene-Lambertz, 1997). In fact, the MMN response is not just larger but rather includes a separate subcomponent when the phoneme boundary is crossed. For example, the same VOT span crossing an English phonemic category boundary evokes a far larger MMN than one that does not (Phillips et al., 2000). These results show that discrete phonological representations can be accessed by the auditory cortex, thus providing the basis for lexical storage and further linguistic computation.

Tonal languages have successfully explored experience-dependent effects on the automatic processing of phonologically contrastive pitch (Xi, Zhang, Shu, Zhang, & Li, 2010; Chandrasekaran, Krishnan, & Gandour, 2007, 2009; Ren, Yang, & Li, 2009; Klein, Zatorre, Milner, & Zhao, 2001; Gandour, Dechongkit, Ponglorpisit, & Khunadom, 1994). Chandrasekaran et al. (2007) showed that early cortical processing of pitch contours might be shaped by the relative saliency of acoustic dimensions underlying the pitch contours of a particular language.

However, very few studies have examined suprasegmental prosodic contrasts that convey discursive or pragmatic meanings in intonational languages, like declarative versus interrogative intent, and their results are controversial. In Doherty, West, Dilley, Shattuck-Hufnagel, and Caplan’s (2004) study, a set of English speakers made judgments about falling statements (e.g., She was serving up the meal), rising declarative questions (with no word order change), and falling questions with the corresponding word order change (e.g., Was she serving up the meal?). The authors found an increased BOLD activity for rising declarative questions over the falling counterparts, and they argued that the differences may reflect the presence of a subtle aspect of illocutionary force (conduciveness) in the utterances with rising intonational contours. Fournier et al. (2010) examined the processing of lexical–tonal and intonational contrasts by speakers of an intonational language (STD Dutch) and of a tonal dialectal variety of Dutch (Roermond Dutch). They assumed that the brain responses to the stimuli would depend on the subjects’ language experience, but no group differences were found. The authors argued that the expression and recognition of discourse meanings by means of intonation, which is considered universal among languages, were not necessarily realized in an identical way in the human brain. Finally, Leitman, Sehatpour, Shpaner, Foxe, and Javitt (2009) employed two artificial sequential sinusoidal tones corresponding to English declaratives and interrogatives. An “interrogative” DEV block and a “declarative” DEV block were presented, and authors found significant MMN responses in both conditions.

In summary, the representation of segmental and tonal phonological distinctions is found to be evident by means of the MMN, but this is not the case of intonational discourse contrasts. The abovementioned MMN results and its magneto-encephalographic counterpart on intonational discourse contrasts could be interpreted as detections of acoustic changes in the stimuli and remain far from signaling intonationally based phonological distinctions indicating different meanings. Moreover, no previous study has examined the processing of intonational across-category contrasts (e.g., between statements and questions) and within-category contrasts (e.g., between two types of statements or two types of questions). The above-mentioned studies exclusively used minimal pairs as their basic stimuli, and furthermore, they did not show any evidence for language-specific phonological representations or traces for intonational contrasts.

Interestingly, in Catalan, a rising–falling intonational sequence can be perceived as a statement or as a counterexpectational question, depending exclusively on the size of the pitch range interval of the rising movement. The two rising–falling pitch contours consist of a rising movement associated with the stressed syllable, followed by a falling F0 movement associated with the posttonic syllables (see Figure 1 in Methods). The following examples in (1) show two typical discourse contexts in which these intonational configurations could be found. A statement context is shown in (1)a, and a counterexpectational question in (1)b. In both cases, the target word petita [pa.ɾi.tə] (“little,” feminine) is typically produced with a

**Figure 1.** Idealized schema of the pitch manipulation in the noun phrase petita [pa.ɾi.tə] (“little,” feminine). Duration of the segments is shown at the top, and the link between each segment is shown at the bottom. The Hz values at the center of the image represent the final frequencies of the extreme stimuli (Steps 00 and 15).
low tone on the first syllable, a rising/high tone associated with the second (stressed) syllable, and followed by a falling/low tone associated with the third (posttonic syllable). The prosodic difference between (1)a and (1)b lies on the pitch range difference between the low and the high tones, which is expanded in the case of counterexpectational questions.

(1) a. Com la vols, la cullera? What type of spoon do you want?  
   Petita, [sisplau]. [I want a] little [spoon, please].
   b. Jo la vull petita, la cullera Petita? [N’estás segur?] I want a little spoon.  
   [A] little [one]? [Are you sure?]

Borràs-Comes, Vanrell, and Prieto (2010) performed a set of behavioral experiments (identification and congruity tasks) and confirmed that a categorical phonological contrast exists between these two types of rising-falling contours (compressed vs. expanded pitch range) and that they cue a statement and a question interpretation, respectively. These results represented further evidence that pitch range differences can be used to cue intonational distinctions at the phonological level, in line with the findings of other languages (Vanrell, Mascaró, Torres-Tamarit, & Prieto, 2010; Savino & Grice, 2007). In turn, this finding strengthened the idea that pitch range differences can cue phonological distinctions in the intonational grammar of a nontonal language like Catalan (Aguilar, De-la-Mota, & Prieto, 2009), thus expanding the inventory of potential grammatical units in the description of pitch movements.

The goal of this study was to test whether the intonational contrasts differentiating statements and counterexpectational questions in Catalan can elicit specific MMN responses, thus providing electrophysiological evidence in favor of the idea that the auditory cortex supports distinctive linguistic representations at the intonational level. The article presents a behavioral identification experiment (Experiment 1) and an ERP study consisting three oddball blocks with the aim of finding electrophysiological evidence for this categorical distinction (Experiment 2).

**EXPERIMENT 1**

In Experiment 1, subjects participated in an identification task whose goal was to identify each of the two meanings (statement and question) for a set of 16 stimuli in a pitch range continuum. The goal of Experiment 1 was twofold. First, to corroborate the phonological role of pitch range expansion in the interpretation of rising-falling intonational contours in Catalan. Second, to determine the pitch region at which the change in categorization occurs and, thus, select the target stimuli for the MMN oddball experiment. The same set of participants was enrolled in the auditory ERP experiment several weeks later.

**Methods**

**Participants**

Fifteen healthy volunteers (three men, aged 19–42 years, mean age = 22.5 years; one left-handed) with no history of neurological, psychiatric, or hearing impairment and with normal or corrected-to-normal visual acuity participated in the experiment. Subjects reported not having any auditory deficiency, gave informed consent, and received monetary compensation for their participation. The study was approved by the ethics committee of the University of Barcelona, according to the Code of Ethics of the World Medical Association (Declaration of Helsinki). All participants were native speakers of Central Catalan, and musicians were excluded.

**Stimuli**

To generate the auditory stimuli, a native speaker of Catalan (the first author of this study) read natural productions of the noun phrase *petita* [pəˈti.ə] (“little,” feminine) with a statement pitch contour and a counterexpectational question pitch contour, and these utterances served as the source utterances for our stimuli (Figure 1). The original noun phrase utterances were pronounced with a rising-falling contour. This rising movement was of 0.9 semitones for the statement and 9.9 semitones for the question. We then converted each syllables’ curve to a plateau (taking the mean Hz values for each segment) and normalized the absolute pitch of the pretonic and posttonic syllables of the two utterances (to their mean values). Then, we restored the observed differences of 0.9 and 9.9 semitones, respectively. The height of the accented syllable of the question-based stimuli was then adapted to the value of the statement stimulus, and no noticeable differences were observed between the stimuli. After this, we normalized the durations of each syllable to the mean values of the two original utterances. The synthesized continuum was created by modifying the F0 peak height in 16 steps (distance between each one = 0.6 semitones; see Figure 1). The speech manipulation was performed by means of Praat (Boersma & Weenink, 2008). Each stimulus lasted 410 msec. Rising movements were realized as a 100-msec high plateau starting 30 msec after the onset of the accented syllable /t/ and were preceded by a low plateau for the syllable [pa] (102.4 Hz, 100 msec). The posttonic syllable [tə] was realized with a low plateau (94.5 Hz, 180 msec). The pretonic and posttonic F0 levels were maintained invariable in all manipulations. The peak height continuum ranged from 105.3 to 188.6 Hz.

**Procedure**

Stimuli were presented to subjects over headphones, and their amplitude was adjusted to a comfortable level. Subjects were instructed to pay attention to the intonation of the stimuli and decide which interpretation was more
likely for each stimulus by pressing the corresponding computer key, namely “A” for afirmació (“statement”) and “P” for pregunta (“question”).

The task consisted five blocks in which all 16 stimuli in the continuum were presented to the subjects in a randomized order, for a total of 80 stimuli. We thus obtained 1200 responses for Experiment 1 (16 steps × 5 blocks × 15 listeners). The experiment lasted approximately 8 min.

Response frequencies and RT measurements were automatically recorded by means of E-Prime version 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA; www.pst-net.com/). The experiment was set up in such a way that the next stimulus was presented only after a response had been given, yet subjects were instructed to press the button as quickly as they could.

Results
A one-way ANOVA was carried out with the proportion of “counterexpectational question” responses as the dependent variable. The data were first checked for the occurrence of possible outliers on the basis of RT. Of 1200 data points, 84 cases were treated as outliers, that is, those cases where the RTs were at a distance of at least 3 SDs from the overall mean (RTs ≥ 1799 msec). These cases were excluded from the analysis.

Figure 2 shows the identification rate (y axis) for the auditory continuum created (x axis). This rate is defined as the proportion of “question” responses that were given over the total. The identification function presents a classic S-shape, revealing that the lowest six stimuli belong to the category “statement” and the highest five stimuli to “question.” The perceptual shift from one category to another occurs in the range of Stimuli 6–11; a full crossover from 16.92% to 85.92% is achieved between these five central steps.

The analysis revealed a significant main effect of the auditory stimulus [F(15, 1100) = 117.624, p < .001]. Tukey HSD post hoc tests revealed two main homogeneous subsets, namely between Stimuli 0–6 and 11–15, so we can set an area of change of categorization between Stimuli 6 and 11. To calculate the boundary value between the two categories, the set of data points was fitted to a logistic regression using SPSS (SPSS, Inc., Chicago, IL, 2008). Thus, we obtained the boundary value calculated from the “b0” and “b1” values given for the logistic curve using the following formula: boundary = −ln(b0) / ln(b1). Hence, when y = 0.5, then x = 8.65 (the boundary is therefore located between Stimuli 8 and 9).

Figure 2 plots averaged RT responses in milliseconds (y axis) for all stimuli (x axis). RTs were measured from the start of the utterance playback (total length of the utterance = 380 msec). The graph indicates longer RTs for central stimuli, with a clear increase observed for Stimuli 7–9, which coincides with the area of change reported in the identification function. As expected, listeners displayed faster RTs in the identification of within-category exemplars than in exemplars representing the category boundaries.

Results of a univariate ANOVA indicated a statistically significant effect of stimulus type on RT measures [F(15, 1100) = 2.678, p = .001]. Duncan post hoc tests revealed a homogeneous subset between Stimuli 0–6 and 10–15 and another one between Stimuli 5–10. This second subset
between Stimuli 5–10 roughly coincides with the area of change of perceptual categorization found in the identification function.

Our behavioral results, thus, indicate that the variation in pitch range is the main cue that Catalan listeners use to decide between a statement interpretation and a counterexpectational question interpretation. Taken together, the identification and RT results clearly show that the two intonational categories under examination are categorically perceived. These results replicate the findings by Borràs-Comes et al. (2010) for Central Catalan. Experiment 2 will test whether this intonational contrast can be neurophysiologically represented as measured with the MMN.

EXPERIMENT 2

The aim of Experiment 2 was to test whether the intonational contrasts differentiating statements and counterexpectational questions in Catalan can elicit a specific MMN response, thus showing electrophysiological evidence supporting that the auditory cortex supports distinctive linguistic representations at the intonational level. We hypothesize that discrete intonational representations, as well as discrete phonological representations, can be represented through symbolic memory traces in the brain (see Pulvermüller & Shtyrov, 2006).

Methods

Participants

The same sample of 15 Catalan speakers that participated in the first experiment volunteered in the present experiment. A period of 4–9 weeks was elapsed between the two experiments.

Stimuli and Procedure

On the basis of the results of Experiment 1 (i.e., a central area of change of categorization and two tails of within-category variation), four auditory stimuli were selected to be contrasted by pairs in three different oddball blocks (Stimuli 00, 05, 10, and 15). The choice was made according to two criteria: (1) The physical distance in semitones between two stimuli within a pair was kept constant (three semitones), and (2) two stimuli had to be classified as belonging to the “statement” category and two to the “question” category. Thus, all contrasts involved the same physical difference, but the central one (Stimuli 05 and 10) involved a categorical difference as well. The idealized intonational contours of the stimuli used are displayed in Figure 3.

The experiment consisted three oddball blocks presented in random order, with short pauses in between. Each oddball block lasted 21 min approximately and contained 720 STD stimuli and 180 DEV stimuli (80% STD and 20% DEV). STD and DEV stimuli were presented pseudorandomly, with the constraint that a DEV stimulus was preceded by a minimum of two STD stimuli. Whereas the lower pitch stimulus acted as a STD, the higher acted as a DEV, resulting in the following oddball blocks: lower (within-category; Step 00 STD and Step 05 DEV), central (across-category; Step 05 STD and Step 10 DEV), and higher (within-category; Step 10 STD and Step 15 DEV).

All stimuli were presented with a fixed SOA of 1400 msec. The onset of the deviance between a pair of stimuli appeared at the second syllable of the token (120 msec after stimulus onset). The use of occlusive phonemes at the beginning of each syllable allowed us to obtain reliably time-locked ERPs (see Pulvermüller, 2005). Participants sat in a comfortable chair in a sound-attenuated and electrically shielded room. They were instructed to ignore the sounds delivered by headphones and watch a silent movie with subtitles. The amplitude of the stimuli was adjusted to a comfortable level. The total duration of the experiment was approximately 100 min, including the EEG recording preparation.

EEG Recording

The EEG was continuously recorded with frequency limits of 0–138 Hz and digitized at a sampling rate of 512 Hz (EEmagine, ANT Software B.V., Enschede, Netherlands). Ag/AgCl electrodes were used for the EEG acquisition, 35 of which were mounted in a nylon cap (Quik-Cap; Compumedics, Abbotsford, Victoria, Australia) according to the international 10–20 system. Vertical and horizontal EOGs were measured from monopolar electrodes placed,
respectively, below and laterally to the right eye. The ground electrode was located on the chest, and the common reference electrode was attached to the tip of the nose. All impedances were kept below 5 kΩ during the whole recording session.

The continuous EEG was further bandpass-filtered offline between 1 and 20 Hz and cut in epochs of 700-msec duration, including a prestimulus baseline of 100 msec, for each DEV and STD in all three conditions (except for the STD following a DEV stimulus; 180 DEV epochs and 540 STD epochs per condition). Epochs with a signal range exceeding 100 μV at any EEG or EOG channel were excluded from the averages, resulting in a mean of 143 DEV epochs (SD = 20.3; minimum of 94) and 325 STD epochs (SD = 47.4; minimum of 213) after rejection.

MMN difference waveforms were obtained by subtracting the ERPs elicited by STD stimuli from those elicited by DEV stimuli. The MMN peak was determined from the Fz electrode as the largest negative peak in the interval of 200–400 msec (80–280 msec after stimulus onset) for all difference waves and subjects separately. Because MMN peak latencies were not significantly different across conditions, MMN mean amplitudes were derived in an 80-msec time window centered on the mean peak latency of the grand average waveforms for all the three conditions (265–345 msec).

Data Analysis

The presence of a significant MMN elicited to each intonational contrast was analyzed by means of one-sample *t* tests on the MMN amplitude at Fz in each of the three conditions separately. The intonational contrast effects on the MMN peak latencies and mean amplitudes at Fz electrode were evaluated with separate repeated measures ANOVAs, including the factor contrast (lower [within-category], central [across-category], and higher [within-category]). Because MMN inverts its polarity below the Sylvian fissure (ref), another repeated measures ANOVA was conducted to assess the effects on the MMN mean amplitude retrieved at mastoid electrodes, with the factors Channel (M1 and M2) × Contrast (lower, central, and higher). The Greenhouse–Geisser correction was applied when appropriate.

In an attempt to relate the electrophysiological responses with behavioral measures, a bivariate correlation analysis was performed between the MMN mean amplitude and the categorization index (CI) for all subjects as well as for the grand mean data. For these specific analyses, the EEG data were rereferenced to combined mastoids to better assess the power of the effects. We defined the CI as the difference between the categorization scores to each of the two stimuli in a pair, thus resulting in three measures per subject: lower (within-category; Step 05 – Step 00 scores), central (across-category; Step 10 – Step 05 scores), and higher (within-category; Step 15 – Step 10 scores). The higher the CI, the higher the categorical difference a subject made between a pair of stimuli (please note that we have steps of 0.2 CI because each stimulus in Experiment 1 was presented five times to each subject). To further test the significance of the obtained correlation values, we estimated the variability of the correlation statistic (Pearson’s correlation coefficient) with the bootstrap method. Bootstrapping is a resampling method that helps to perform statistical inferences without assuming a known probability distribution for the data. In short, the correlation index was calculated for 10,000 randomly chosen samples (with replacement) of *n* = 45 (15 subjects × 3 conditions) of MMN amplitude values and CI scores, respectively. The obtained distribution (H1; centered at the Pearson’s coefficient value that is obtained performing a simple correlation with the raw data) was tested for significance against the null hypothesis distribution (H0), which arises from performing the correlation analysis in 10,000 random samples of MMN and CI scores (*n* = 45) pooled together. Thus, the bootstrap method yields a mean of the correlation statistic for the H0 centered at 0, with the confidence intervals (95%) that are used to test the significance of the obtained H1.

Results and Discussion

Grand average waveforms elicited to STD (dotted line) and DEV (continuous line) stimuli at Fz, M1, and M2 electrodes are shown in Figure 4. DEV minus STD stimuli difference waveforms are shown in Figure 5. The mean values of the DEV minus STD waveforms at the 266–346 msec window (and their SDs) are shown in Table 1. The amplitude enhancement of the DEV stimuli AEPs compared with the STD stimuli ERPs, around 180-msec postdeviance onset and identified as the MMN, was statistically significant in each intonational contrast (lower [within-category] contrast, *t* = −6.217, *p* < .0005; central [across-category] contrast, *t* = −8.875, *p* < 10−5; higher [within-category] contrast, *t* = −6.551, *p* < .0005). A repeated measures ANOVA on the MMN peak latencies did not yield any difference between the three conditions (*F*2, 28 = 2.828, *p* = ns, η² = .168). As we hypothesized, the mean amplitude of the MMN was larger for the central (across-category) intonational contrast (Steps 05–10) compared with the within-category contrasts: intonational contrast effect at Fz, *F*2, 28 = 3.417, *p* < .05, η² = .196 (within-subject contrasts: lower vs. central, *F*1, 14 = 6.256, *p* < .05, η² = .309; central vs. higher, *F*1, 14 = 4.898, *p* < .05, η² = .259; lower vs. higher, *F*1, 14 = 0.172, *p* = ns, η² = .012). The analysis at the mastoid electrodes yielded similar results to those obtained at Fz, *F*2, 28 = 6.978, *p* = .01, η² = .335 (within-subject contrasts: lower vs. central, *F*1, 14 = 43.403, *p* < .00001, η² = .756; central vs. higher, *F*1, 14 = 4.323, *p* = .056, η² = .256; lower vs. higher, *F*1, 14 = 1.203, *p* = ns, η² = .079). The scalp distribution maps of the MMN are shown in Figure 6.

Furthermore, an analysis between the CI and the MMN mean amplitude (electrophysiological measure) yielded a
significant negative correlation, Pearson’s correlation statistic $= -0.308, p < .05$ (one-tailed). This means that the higher the amplitude of the MMN elicited in an oddball sequence with that pair of stimuli acting as DEV and STD stimuli, the more a subject categorized differently the two stimuli within a pair. The significance of this correlation was further supported by an analysis using the bootstrap method: Pearson’s correlation statistic sampling distribution centered at $-0.308$, confidence interval of the null hypothesis with $95\%$ confidence bounds $[-0.289, 0.297], p = .018$. Additionally, we performed a bivariate correlation between the grand mean of the CI and the grand mean of the MMN, yielding a significant Pearson’s correlation of $-0.999, p = .011$. We acknowledge that the statistics on the grand mean cannot be taken as a real proof of the existence of a correlation between the CI and the MMN; however, it serves us to illustrate more clearly the direction of the effects. Bivariate correlations between CI and MMN for all subjects and grand means, respectively, and the bootstrap sampling distributions of the alternative and null hypotheses can be seen in Figure 7.

**GENERAL DISCUSSION**

Previous electrophysiological studies on vocalic and consonantal phonological contrasts have found evidence
that native linguistic contrasts elicit significantly larger 
MMN responses than nonnative contrasts (Eulitz & Lahiri, 
2004; Winkler et al., 1999; Näätänen et al., 1997). In ad-
dition, acoustic contrasts that cross a category boundary 
lead to larger MMN responses than comparable acoustic 
contrasts that did not cross these category boundaries 
(Phillips et al., 2000; Sharma & Dorman, 2000; Dehaene-
Lambertz, 1997). Similarly, it is an established result that 
tone contrasts in tonal languages obtain larger MMN re-
sponses when listeners are exposed to native tonal con-
trasts (Chandrasekaran et al., 2009; Ren et al., 2009; Klein 
et al., 2001; Gandour et al., 1994) and also in tonal stim-
uli crossing the category boundaries (Xi et al., 2010; 
Chandrasekaran et al., 2007). Thus, a substantial set of em-
pirical results demonstrates the larger activation of mem-
ory traces for linguistic elements in the human brain. In 
line with this, Näätänen (2001) proposed that the MMN re-
flects the early access to stored linguistic representations. 
In the recent years, more evidence has been accumulat-
ing that MMN reflects the early access of linguistic infor-
mativ information has been ignored” (Leitman et al., 2009, 
p. 289). Fournier et al. (2010) argued that the recognition 
of discourse meanings by means of intonation was not 
necessarily clear by looking at the human brain. 

Our results go beyond the body of evidence presented 
by previous experiments and provide electrophysiological 
evidence that phonological contrasts at the intonational 
level (based on a pitch range difference) are encoded in 
the auditory cortex. The empirical data in our study were 
based on an intonational contrast between statements and 
questions in Catalan. The results of Experiment 1, which 
tested the participants’ interpretation of isolated stimuli 
in a binary way (statement vs. counterexpectational ques-
tions), corroborated the findings of Borràs-Comes et al. 
(2010) by indicating a clear nonmonotonic identification. 
Specifically, a perceptual shift from one category to an-
other occurred in the range of Stimuli 6–11, with a full 
crossover from 16.92% to 85.92% achieved between these 
five central steps. Moreover, post hoc tests revealed two 
main homogeneous subsets, namely between Stimuli 0– 
6 and 11–15. Concerning RTs, listeners displayed faster 
RTs in identification of within-category exemplars than in 
exemplars representing the category boundaries (espe-
cially for Stimuli 7–9). For Experiment 2, four auditory 
stimuli were selected to be contrasted by pairs in three dif-
derent oddball blocks. Although the physical distance be-
tween each pair of stimuli was kept constant, the central 
pair represented an across-category contrast, whereas the 
other pairs represented within-category contrasts. The 
mean amplitude of the MMN was found to be larger for 
the across-category contrast compared with the other con-
trasts, suggesting that intonational contrasts in the target 
language can be encoded automatically in the auditory 
cortex. Moreover, our results showed that the activation 
of these auditory cortex intonational representations was 
related to the individuals’ subjective perception and per-
formance. As Pulvermüller and Shtyrov (2006) proposed, 
the MMN might reflect not only the automatic detection 
of a change but also the activation of a certain symbolic 
memory trace in the brain. Finding an MMN for within-
category contrasts would indicate that a change in the 
acoustic environment has been detected, but the symboli
memory trace is still the same called by the STD. By contrast, finding a significantly larger MMN in an across-category contrast would thus indicate not only a reactivation of the attentional system but also an activation of different cortical cell assemblies supporting another long-term memory trace.

It is also important to note that our data can also support an alternative explanation, that is, that the MMN results may reflect perceptual saliencies or distinctiveness that may be consistent across languages. Although external evidence suggests that the MMN may reflect symbolic memory traces, others have suggested that the MMN robustness may reflect individual differences in dimensional weighting (e.g., Chandrasekaran et al., 2007, 2009). For example, animals show categorical perception (Kuhl & Miller, 1978), and thus, the increased MMN for across-category contrasts may reflect auditory discontinuities (e.g., Holt, Lotto, & Diehl, 2004, for VOT), that is, natural boundaries within which distinctiveness is enhanced, reflecting a warped acoustic space (Kuhl & Miller, 1975). One possibility for demonstrating the explanation based on symbolic memory traces would be the application of a cross-language design, but this should be addressed in future studies.

The present experiment design does not allow us to draw any conclusions regarding the specific neural network supporting the across-category intonation contrasts observed here as enhanced MMNs, and therefore, we can only speculate. The MMN has multiple cerebral sources, including the auditory cortex (Escera, Alho, Schröger, & Winkler, 2000; Alho, 1995) and frontal regions (Deouell, 2007), and recent results from animal (Antunes, Nelken, Covey, & Malmierca, 2010; Malmierca, Cristaudo, Perez-Gonzalez, & Covey, 2005; Ulanovsky, Las, & Nelken, 2003) and human studies (Grimm, Escera, Slabu, & Costa-Faidella, 2011; Slabu, Escera, Grimm, & Costa-Faidella, 2010) have suggested that deviance detection yielding to MMN generation might encompass the whole auditory hierarchy (Grimm & Escera, in press). Moreover, recent studies have suggested that processing linguistic DEV features recruits not only auditory but also motor cortical regions in a somatotopic fashion (Hauk, Shtyrov, & Pulvermüller, 2006; Shtyrov, Hauk, & Pulvermüller, 2004) and that category-based enhancement is often found in prefrontal regions (Freedman, Riesenhuber, Poggio, & Miller, 2001). In addition, Raizada and Poldrack (2007) found that lower-level auditory areas show little enhancement of across-category phonemic pairs relative to higher-order areas, and Zhang et al. (2011) have shown that across-category variation on a lexical tonal continuum activated the left middle temporal gyrus, apparently reflecting abstract phonological representations, whereas the within-category contrasts activated the superior temporal and Heschl gyri bilaterally. Therefore, it is possible that the cross-category intonational effects observed here as a frontally distributed enhanced MMN, compared with the within-category one, might reflect the activation of a distributed cortical network including higher-order auditory areas, such as the posterior superior temporal and middle temporal gyri and frontal regions.

In summary, the MMN findings reported in this study show that a distributed auditory-frontal cortical network supports phonological representations not only at the segmental level but also at the intonational level. Catalan listeners showed a larger MMN response to differences in pitch, activating the semantic contrast between a question and a statement. To our knowledge, this is the first study showing a clear electrophysiological response to a change of intonational category. This result agrees with Pulvermüller and Shtyrov’s (2006) hypothesis that MMN responses reflect early automatic processes not only affecting lexical access and selection but also semantic and discourse information processing.

Acknowledgments
This study was supported by the Consolider-Ingenio 2010 (CSD2007-00012) Program, the grant PSI2009-08063 of the Spanish Ministry of Science and Innovation, and the Generalitat de Catalunya (SGR2009-11). C. E. holds the ICREA Academia Distinguished Professorship.
REFERENCES


Niätänien, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology, 38*, 1–21.


AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

During the preparation of your manuscript, the questions listed below arose. Kindly supply the necessary information.

1. Please check affiliation of Pilar Prieto.
2. Please update if Grimm & Escera (in press) has already been published.

END OF ALL QUERIES