cortex xxx (2010) 1–12

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cortex

Research report

Detection of pitch violations depends upon the familiarity of intonational contour of sentences

Lucia Colombo [a,*], *Chizuru Deguchi* [a], *Magali Boureux* [a], *Michela Sarlo* [a] *and Mireille Besson* [b]

[a] *Dipartimento di Psicologia Generale, University of Padua, Italy*
[b] *Institut de Neurosciences Cognitives de la Méditerranée CNRS – Université de la Méditerranée, Marseille, France*

ARTICLE INFO

Article history:
Received 2 October 2009
Reviewed 11 January 2010
Revised 11 March 2010
Accepted 30 March 2010
Action editor Roberto Cubelli
Available online xxx

Keywords:
Pitch detection
Prosody
Intonation
ERPs
Jabberwocky sentences

ABSTRACT

Introduction: Former studies have shown that a pitch change in utterances in speech was detected accurately in both native (French) and unfamiliar (Portuguese) language and produced an early negativity and a late positivity in the event-related brain potentials (ERPs), more clearly marked in the native language (Schön et al., 2004; Marques et al., 2007). The present study used the same design to further investigate the influence of the familiarity of the language context on pitch perception with Italian participants. The aim was to examine the effects of pitch change in the native (Italian) and foreign (French) languages, and in meaningless sentences preserving the intonational contour of the mother tongue (jabberwocky).

Method: Weak and strong pitch changes incongruous with the preceding context were compared to a control congruous condition. Participants had to decide as fast and as accurately as possible if they perceived a pitch change. Both behavioral (accuracy and reaction times – RTs) and ERP measures were analyzed.

Results: Results showed optimal accuracy and faster RTs in Italian, followed by jabberwocky, and then French. The same trend was present in ERP data, with an early negativity over temporal sites and a late positivity over parietal sites. These effects developed earlier and were more pronounced for the strong incongruity in Italian and in jabberwocky than in French.

Conclusions: The similarity of results for Italian (congruous) and jabberwocky sentences on one hand, and the difference of results for French sentences, on the other hand, show that familiarity with intonational contour of utterances/speech provided essential cues to perform the task.

1. Introduction

Prosodic information is very important for speech comprehension and interpretation, because it conveys linguistic and emotional information. Linguistically, it is highly correlated with lexical, morphological, semantic, syntactic, and pragmatic speech structures. For example, the relationship between the syntactic structure of sentences and prosodic information has been largely investigated. When syntactic and prosodic interpretations of a spoken sentence are inconsistent,

* *Corresponding author*. Dipartimento di Psicologia Generale, Via Venezia 8, 35131 Padova, Italy.
E-mail address: lucia.colombo@unipd.it (L. Colombo).
0010-9452/$ – see front matter © 2010 Elsevier Srl. All rights reserved.
doi:10.1016/j.cortex.2010.03.008

processing difficulties ensue (D'Imperio et al., 2007; Marslen-Wilson et al., 1992; Warren et al., 1995). Moreover, listeners exploit prosodic information to resolve local and global structural ambiguities in syntactic interpretation (Christophe et al., 2003, 2004; for a review, see Cutler et al., 1997). Studies using event-related brain potentials (ERPs) have exploited mismatch in syntax and prosody to investigate the time course of prosodic processing. Results (Pannekamp et al., 2005; Steinhauer et al., 1999; Steinhauer and Friederici, 2001; Steinhauer, 2003) showed that a positive component (closure positive shift – CPS) with maximum amplitude around 300 msec was elicited by clause boundaries, thereby indicating that listeners exploit prosodic information to assign structural relations. The CPS probably belongs to the P3-family of components and is functionally similar to the P3b component elicited by task-relevant and unexpected events. Moreover, by using an orthogonal design, Astésano et al. (2004) were able to demonstrate interactive effects of prosodic processing (i.e., processing of declarative and interrogative sentences, also reflected by a delayed positivity, the P800, probably belonging to the P3-family) and semantic processing (reflected by an N400 component). However, prosodic contours that are unexpected, infrequent or incongruous with other sentence-level information are also reflected in other ERP components (N1, N2, P3a,...) depending on position of stimuli in the sentence, task demands and position of incongruities in the sentence (Eckstein and Friederici, 2006; Heim and Alter, 2006; Magne et al., 2006, 2005; Mietz et al., 2008). In general, these results suggest that expectancies based on sentence interpretation at different levels (lexical, semantic, syntactic and pragmatic) are generated either automatically or intentionally. Early components are most often considered as reflecting automatic expectancies (Steinhauer et al., 1999) and later components as reflecting strategic and more controlled processes, like rehearsal, analysis, integration and decision (Friedrich et al., 2001; Hahne and Friederici, 1999; Isel et al., 2005).

Prosodic or suprasegmental information can be described at the acoustic level by considering fundamental frequency or pitch (F0), intensity, duration, and spectral characteristics. Differences in F0 provide information to detect pitch contour (D'Imperio, 2002; Halliday and Greaves, 2006), and are important cues to locate lexical stress (Morton and Jassem, 1965). Detection of incongruities in pitch contour (i.e., strong and weak pitch changes) was investigated in recent studies by Besson et al. carried out on French adults (Schön et al., 2004) and children (Magne et al., 2006) listening to their native language or to a foreign language, Portuguese (Marques et al., 2007). In order to verify the hypothesis that musical training would result in higher pitch discrimination ability, participants were either expert musicians or non-musicians. Both were faster and more accurate to discriminate strong compared to weak pitch changes, defined on the basis of the percentage increase compared to the control condition (i.e., 200% or 135%, respectively). Moreover, the amplitudes of an early negativity (peaking around 150 msec) and of a later, more prominent positive component were larger to strong than to weak incongruities (Schön et al., 2004).

Overall, this pattern of results was reproduced even when French participants, musicians and non-musicians, perceived F0 changes in Portuguese sentences, a language with which they were not familiar (Marques et al., 2007), with a reliable effect of musical expertise. However, there were also some interesting differences. First, reaction times (RTs) to strong pitch changes in the foreign language were slower, and error percentage was higher compared to the native language. In addition, the onset of the positive component to the strong incongruity was delayed (by about 300 msec) in Portuguese compared to French. Finally, although a negative component was present, it did not show the same morphology and time course as in Schön et al.'s study (2004). Marques et al. (2007) suggested that RTs and ERP components were delayed, since participants did not understand the language they were listening to, and were not able to develop expectancies regarding when the sentence-final word would occur.

When the intonational contour of a sentence is disrupted by a sudden change in F0, different processes are involved, with probably automatic perceptual processing of the F0 variation (detection of the disruption) followed by less automatic processes to analyze the change in pitch, to verify its incongruity within the sentence prosodic contour, and to categorize it in order to provide a response. These computations can be carried out more easily when the prosodic contour of the sentence is familiar (i.e., in the native language), but they also take place in a foreign language, with an unfamiliar intonational pattern. Marques et al. (2007) suggested that the ability to detect prosodic incongruities in a sentence context might reflect the interaction between semantic and prosodic processes. The RTs and ERPs latency differences between French and Portuguese would mainly arise because participants did not understand Portuguese and were consequently unable to predict when the critical final word would be presented.

However, an alternative interpretation is that detection of pitch changes is largely based on the prosodic contour of the sentence. Presumably, speakers implicitly acquire typical sentence intonational contours in their native language, and this implicit learning, together with syntactic and semantic information, allows a listener to predict the prosodic suprasegmental patterns. In turn, this information can at least partially help segmenting the speech input into words and phrases.

There is evidence that listeners exploit distal prosodic cues in spoken language processing (i.e., cues that are far from the sentence position where segmentation occurs). For example, relative size of prosodic boundaries affects the way an ambiguous syntactic structure is parsed (Schafer et al., 2000). Dilley and McAuley (2008) investigated how, based on auditory perceptual organization principles, listeners exploited F0, duration or both as distal cues to define lexical boundaries and to segment speech. Pannekamp et al. (2005) showed that the CPS occurred not only with normal sentences, but also with jabberwocky sentences (i.e., pseudowords replacing content words) or hummed sentences, in which only prosodic information was present. Thus, it is likely that prosodic structures of the native language are implicitly acquired and form distal cues that are used to organize syntactic structures and lexical boundaries. Recent studies provide evidence that speech perception is tuned to language experience (Kuhl et al., 2006), and that adults tend to perceive speech sounds according to the phonological systems of their native language (Best and

Please cite this article in press as: Colombo L, et al., Detection of pitch violations depends upon the familiarity of intonational contour of sentences, Cortex (2010), doi:10.1016/j.cortex.2010.03.008

Strange, 1992; Flege, 1995). However, familiarity with a non-native language can also aid in exploiting phonetic cues even though they are not part of the native phonetic inventory (Cho and McQueen, 2006).

In the present study, the familiarity of distal prosodic information was manipulated in order to determine whether discrimination of pitch changes (strong or weak) depends on the possibility to carry out a syntactic and semantic interpretation of the sentence context or may solely rely on the familiarity and predictability of the intonational contour of the context. We used an experimental design similar to Schön et al. (2004) and Marques et al. (2007) by presenting to Italian participants sentences in which F0 of whole final words was slightly or strongly increased. Sentences were spoken in Italian (native language), French (foreign unknown language) or in jabberwocky sentences. The Italian–French contrast extends the French–Portuguese comparison investigated in separate studies by Schön et al. (2004), and by Marques et al. (2007) by using a within-subject design. Both Italian and French are syllable-timed languages (i.e., languages in which syllabic duration tends to be similar Cutler and Norris, 1988), as opposed to stress-timed languages, like English, where it is different. However, Italian is characterized by lexical stress, whose position is variable and unpredictable, whereas French has a single rhythmic stress that is assigned to the final full syllable (i.e., not containing a schwa) of the last lexical item of a stress group (Di Cristo and Hirst, 1998; Magne et al., 2007). In French, there is a reciprocal relationship between accent and intonation, since there is a pitch change on the last syllable (the accented one) of the rhythmic group (i.e., a group of syntactically-related words, like prepositional or verbal phrases). In Italian, word-stress increases the duration of the tonic syllable but it isn't necessarily marked by pitch variation. This only occurs when stress is located on the final syllable of a group. Consequently, variations in pitch would not be confused with stress marking and could only be interpreted to signal intonational changes (Rossi, 1998).

While Italian sentences included all different types of information (prosodic, semantic and syntactic), jabberwocky sentences preserved the familiar prosodic intonation of Italian but with only limited syntactic information, due to function words and morphological affixes, and no semantic content. If Italian speakers are sensitive to the intonational patterns of their own language, they might use them to detect unacceptable pitch changes. Moreover, familiar intonational patterns might also help participants to predict when the sentence-final word might occur. Thus, if pitch incongruity discrimination is based on the intonational contour of the sentence, the level of performance should not differ between Italian and jabberwocky sentences (both contain familiar prosodic information), and should be better than with French sentences (with unfamiliar prosodic contour). Moreover, the similarity of the ERP pattern should be higher between Italian and jabberwocky than between Italian and French. Specifically, we expected an early negativity to be elicited by strong incongruity in both Italian and jabberwocky (as was found for the mother tongue by Schön et al., 2004) but not in French (as was found for a foreign language by Marques et al., 2007). Moreover, we hypothesized that the amplitude of the late positivity (P3-like component) reported in previous studies (e.g., Marques et al., 2007; Schön et al., 2004) should be larger for both Italian and jabberwocky sentences than for French sentences. By contrast, if a semantic interpretation of the sentence is crucial in order to identify the word on which pitch change occurs, as proposed by Marques et al. (2007), and if participants are not able to exploit familiar prosodic information without semantic content, the level of performance for Italian sentences should be higher than for both jabberwocky and French sentences, that should not differ. At the ERP level, the amplitude of the P3-like component should be similar for both French and jabberwocky sentences and smaller than for Italian sentences. The contrast between Italian, French and jabberwocky should consequently allow us to verify if pitch incongruity detection is based on prosody, on semantics or on both.

2. Experiment

2.1. *Method*

2.1.1. *Participants*

The study was conducted in compliance with the ethical standards of the Declaration of Helsinki and of the local Ethical Committee. Thirty-nine Italian native speakers participated in the experiment. Because of the difficulty to discriminate the weak incongruity condition in French, the level of accuracy for 9 participants was too low, and their data were removed from further analyses. Moreover, one participant had too many artifacts in the ERP recordings and only behavioral data were included. Mean age of the 30 participants was 24.77 years ($SD = 4.17$). All participants were right-handed and had normal hearing according to self-report. None of them had learned or was able to understand French. They were paid for participation in the experiment that lasted 2 hours and a half.

2.1.2. *Materials*

Ninety-nine sentences were created for each language condition (Italian, jabberwocky and French). French and Italian sentences were declarative sentences with one or two main clauses, and/or a subordinate or a prepositional phrase. French sentences were partially taken from the materials used in Schön et al. (2004), but were mostly new sentences, with the constraint that they ended with a disyllabic word starting with a stop consonant, as Italian and jabberwocky stimuli did. Last words of Italian sentences were stressed on the antepenultimate syllable (as the majority of Italian words), while those of French sentences were stressed on the last syllable (as all French words are). Jabberwocky sentences were created by substituting the content words of the Italian sentences with pseudowords with the same number of syllables. All sentence types were matched for number of syllables, and as much as possible for syntactic structure. Examples of the sentences are the following:

Italian: Marco non deve andare in bicicletta sulla corsia riservata ai taxi.
Jabberwocky: Forvo non dusna fadare in chirinetta sulla gornea tirubata ai tessi.

French: Le père de Sophie est accueilli à son arrivée par un délicieux gratin.

The 297 sentences were spoken by a fluent Italian–French bilingual female speaker, recorded and digitized using the computer software "Audacity". Italian and French sentences were read with a natural intonation appropriate to each language. Each jabberwocky sentence was read immediately after the corresponding Italian sentence, so that the speaker maintained a similar prosodic contour in both conditions. Mean sentence duration was 3909 msec (SD = 509) for Italian, 4393 msec (SD = 612) for jabberwocky and 4104 msec (SD = 537) for French. Mean duration of the final word was 507 msec (SD = 60) for Italian, 535 msec (SD = 61) for jabberwocky, and 511 msec (SD = 127) for French.

Three lists were built so that each sentence was presented in each of the three experimental conditions across participants. The final words of the sentences were prosodically congruous, weakly incongruous, or strongly incongruous with the preceding context. To create the incongruous conditions, Italian and French sentences were modified by using the software "Praat" (Boersma and Weenink, 1996). F0 of the whole last word was increased to 115% for the weak incongruity, and to 200% for the strong incongruity conditions, while maintaining the same pitch contour. The increase in F0 for the weak incongruity was based on the results of pretests, so that accuracy levels in Italian would be comparable to that obtained by Schön et al. (2004) in French.

Thus, a total of 891 sentences was used in the experiment (99 sentences × 3 languages × 3 congruity conditions), with 33 sentences in each experimental condition. Each participant received one list of 297 sentences (33 trials × 3 languages × 3 congruity conditions). Each sentence was only presented in one condition.

2.1.3. *Procedure*

The experiment took place in an electrically shielded quiet room, and was programmed and presented using the software "e-prime". Participants were seated in front of a PC monitor and were required to fixate the central asterisk while attentively listening to the sentences that were auditorily presented through a pair of speakers symmetrically positioned in front of them. They were instructed to decide as quickly and accurately as possible, whether the intonation of the final word of each sentence was congruous or incongruous with the former sentence context, pressing one of two buttons of a response box (one for "congruous" and the other for either weakly or strongly incongruous). RTs were measured from the onset of the final word (corresponding to the onset of the F0 modification in the incongruous conditions) until a response button was pressed.

For each participant, the three language conditions were presented separately in three experimental blocks. Each experimental block began with a practice block of 9 trials, during which feedback on the correctness of the response was given, to familiarize participants with the task and to train them to blink during the inter-trial interval. Feedback was only given on practice trials. In each block, sentences were presented in a random order. The hand of response and the order of presentation of the three language conditions were counterbalanced across participants.

2.1.4. *Electroencephalogram (EEG) recordings and analyses*

EEG was recorded from 31 scalp electrodes, mounted on an elastic cap, and located at the following sites according to International 10/20 system: Fpz, Fz, FCz, Cz, CPz, Pz, Oz, Fp1, Fp2, F3, F4, FC3, FC4, C3, C4, CP3, CP4, P3, P4, O1, O2, F7, F8, FT7, FT8, T3, T4, TP7, TP8, T5, T6. These recording sites plus an electrode placed on the right mastoid were referenced online to the left mastoid electrode and digitally re-referenced offline to the algebraic average of the left and right mastoids. Electrode impedances never exceeded 9 kΩ. For the purpose of artifact scoring, vertical and horizontal electro-oculograms (EOGs) were recorded. Electrode pairs (bipolar) were placed at the supra- and sub-orbit of the right eye and at the external canthi of the eyes. The EEG and EOG signals were amplified by a Neuroscan Synamp amplifier (El Paso, TX, USA) with a bandpass filter of .05–100 Hz and were digitized at 500 Hz (16 bit AD converter, accuracy .084 μV/LSB) and stored on a Pentium II computer. The notch filter was set at 50 Hz.

Continuous EEG data were corrected for eye blinks using a regression-based correction algorithm (Scan 4.1 software). EEG was then segmented off-line into single epochs from 150 msec before to 1600 msec after the onset of the sentence-final word. EEG epochs were baseline-corrected against the mean voltage during the 150-msec pre-stimulus period. All EEG epochs were visually scored for eye movement and other artifacts, and each portion of data containing artifacts greater than ±70 μV in any channel was rejected for all the recorded channels prior to further analysis. Artifact-free trials with correct behavioral responses were separately averaged for each subject in each experimental condition.

ERP data were analyzed by computing mean amplitudes, relative to a 150-msec baseline. Latency windows were selected based upon visual inspection of the traces and on preliminary 100 msec latency band analysis between 0 msec and 1200 msec after the onset of the sentence-final word. When the effects were similar in successive 100-msec windows, latency bands were pooled together. For each language, ANOVAs were computed for midline and lateral electrodes separately and included Language (Italian, jabberwocky and French), Congruity (congruous, weakly incongruous, and strongly incongruous) and Electrodes (Fz, Cz and Pz) as within-subjects factors for midline analysis; and Language (3), Congruity (3), Hemisphere (2), ROI (3 regions of interest: fronto-central, temporal, and parietal) and Electrode (3) as within-subjects factors for lateral analysis. In each ROI three electrodes were grouped together (F3, F7, FC3 and F4, F8, FC4; C3, T3, TP7 and C4, T4, TP8; CP3, P3, T5 and CP4, P4, T6). All p values were adjusted with the Greenhouse–Geisser epsilon correction for nonsphericity. Tukey tests were used for all post-hoc comparisons.

3. Results

3.1. *Behavioral data*

RTs data were analyzed for correct responses only. ANOVAs were carried out on both accuracy and RTs, with congruency (congruous, weak incongruity, strong incongruity) and language (Italian, jabberwocky and French) as within-subject factors.

Please cite this article in press as: Colombo L, et al., Detection of pitch violations depends upon the familiarity of intonational contour of sentences, Cortex (2010), doi:10.1016/j.cortex.2010.03.008

3.1.1. Accuracy

Mean RTs and accuracy rates for the two groups of participants are displayed in Fig. 1. As accuracy was at ceiling in the strong incongruity condition, statistical analyses were only carried out on the weakly incongruous and congruous conditions. Results showed a main effect of Language [$F(2,58) = 26.36$, MSE = .01, $p < .001$]. Post-hoc tests showed that the level of performance was similar for Italian (.77) and jabberwocky (.75) and lower for French sentences (.65). The main effect of Congruity was also significant [$F(1,29) = 125.24$, MSE = .06, $p < .001$] with higher accuracy in the congruous condition (.93) than in the weakly incongruous condition (.52). The Language by Congruity interaction was significant [$F(2,58) = 7.12$, MSE = .01, $p < .01$]. In the congruous condition, accuracy was higher for Italian than for French sentences [.95 *vs* .90; $t(29) = 2.90$, SE = .02, $p < .01$]. Performance on jabberwocky (.93) did not differ from Italian, and only marginally from French [$t(29) = 1.81$, SE = .02, $p = .08$]. In the weak incongruity condition the difference between Italian (.60) and French sentences (.41) was also significant [$t(29) = 5.15$, SE = .04, $p < .001$] but the difference between jabberwocky (.57) and French sentences was significant [$t(29) = 4.42$, SE = .001]; again the difference between Italian and jabberwocky sentences was not significant.

3.1.2. RTs

Results showed main effects of Language [$F(2,58) = 22.31$, MSE = 18,505.29, $p < .001$] and Congruity [$F(2,58) = 175.13$, MSE = 22,098.44, $p < .001$] and a significant interaction [$F(4,116) = 3.48$, MSE = 7054.91, $p = .01$]. Inspection of Fig. 1b and Table 1 shows that the Italian–jabberwocky difference was not significant in the congruous and weakly incongruous conditions, but it was reliable in the strongly incongruous condition. In contrast, French sentences always significantly differed from both Italian and jabberwocky sentences except in the strong incongruity, where French was not slower than jabberwocky.

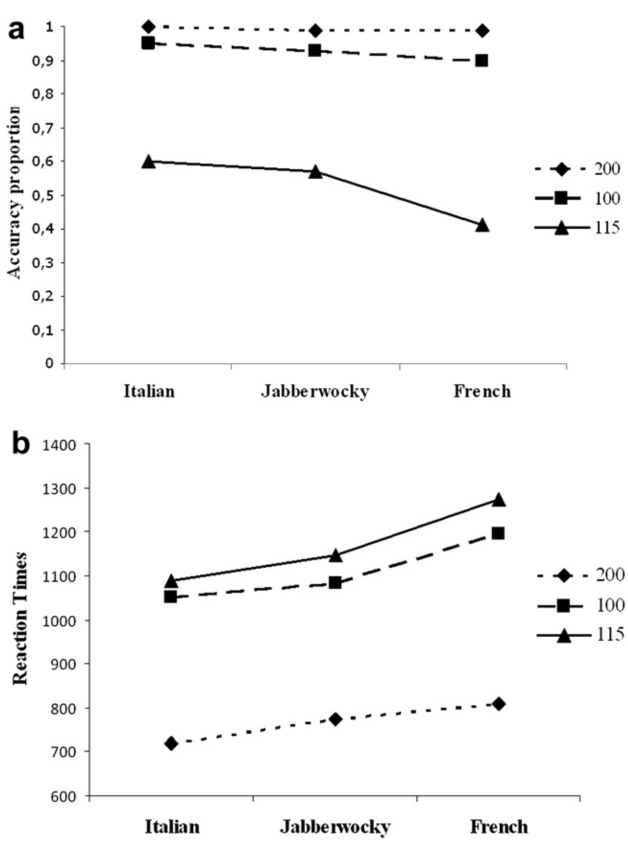

Fig. 1 – Mean accuracy proportion (top) and RTs (bottom) in the weak incongruity, strong incongruity and congruous conditions for Italian, jabberwocky and French.

3.2. EEG analysis

Results of the general ANOVAs are reported in Table 2. In the 0–200 msec latency range and at lateral electrodes mean amplitudes were more negative for Italian (−1.10 μV, $p < .001$) and jabberwocky (−.89 μV, $p < .01$) than for French (−.44 μV; see Table 2). Moreover, in the 100–200 msec range, an early negative component was larger for strong incongruity (midline = −1.98 μV and lateral = −2.08 μV) than for weak incongruity (midline = −.86 μV and lateral = −.88 μV, $p < .05$ and $p < .001$, respectively) and congruous endings (midline = −.61 μV and lateral = −.75 μV, $p < .01$ and $p < .001$, respectively). The Language by Congruity interaction was also significant at midline and lateral electrodes in both the 200–400 msec and the 500–700 msec latency bands.

To further analyze these interactions, separate ANOVAs were conducted for each congruity condition, including as factors the three sentence conditions (Italian, jabberwocky and French) as well as Electrodes (Fz, Cz and Pz) for midline analysis and Hemispheres, ROIs and Electrodes (as defined above) for lateral analysis. Results are presented on Table 3. For strong incongruity and at both midline and lateral electrodes, mean amplitudes in the 0–100 msec range (N1 component) were smaller for French (midline = −.09 μV and lateral = −.02 μV; see Fig. 2) than for Italian (midline = −.61 μV and lateral = −.53 μV, $p < .05$ in both conditions) and jabberwocky (midline = −.72 μV and lateral = −.58 μV, $p < .05$ and $p < .01$, respectively). Moreover, in the 200–500 msec range and at parietal sites the positivity (P3b component) was larger for Italian (midline = 8.34 μV and lateral = 5.90 μV), than for jabberwocky (midline = 6.61 and lateral = 4.43 μV, $p < .001$ in both conditions) and for jabberwocky compared to French (midline = 4.76 μV and lateral = 3.07 μV, $p < .001$ in both

Table 1 – Results of *t*-tests (*t*) on RTs data and standard errors (SE) comparing Italian and jabberwocky sentences (It–Jab), Italian and French sentences (It–Fr) and French and jabberwocky sentences (Fr–Jab) in each congruity condition.

Contrasts	Pitch variations					
	Congruous		Weak		Strong	
	t	SE	t	SE	t	SE
It–Jab	<1	22.5	<1	29.4	2.3*	18.2
It–Fr	5.2**	26.8	5.1**	33.9	4.5**	18.5
Fr–Jab	5.2**	24.8	2.9*	38.8	<1	22.2

*$p < .05$; **$p < .001$.

Please cite this article in press as: Colombo L, et al., Detection of pitch violations depends upon the familiarity of intonational contour of sentences, Cortex (2010), doi:10.1016/j.cortex.2010.03.008

Table 2 – Results of overall ANOVAs on mean amplitudes of ERPs at midline and lateral electrodes.

		Latency band (msec)	Factors	*df*	*F*	*p*
Overall data	Midline	100–200	C	2, 56	6.41	*
		200–400	L × C	4, 112	8.55	**
		500–700	L × C	4, 112	6.01	**
	Lateral	0–200	L	2, 56	11.43	**
		100–200	C	2, 56	12.98	**
		200–400	L × C	4, 112	9.21	**
		500–700	L × C	4, 112	7.35	**

L: language (Italian *vs* jabberwocky *vs* French); C: congruity (congruous *vs* weak incongruity *vs* strong incongruity).
df: Uncorrected degrees of freedom; *p*: probability level after Greenhouse–Geisser correction of *F*-values.
**: $p < .001$; *: $p < .01$.

conditions). No differences were found between 500 and 700 msec. Finally, the Language by Electrode and Language by ROI interactions were also significant in the 700–900 msec range. A late positivity developed at parietal sites, that was larger for French (midline = 7.13 μV and lateral = 5.42 μV) than for jabberwocky (midline = 5.74 μV and lateral = 4.43 μV, $p = .001$ and $p < .001$, respectively) and intermediate for Italian (midline = 6.33 μV and lateral = 5.20 μV), which differed only from jabberwocky ($p < .05$). By contrast, at frontal sites, it was larger for Italian (midline = 2.37 μV and lateral = 1.03 μV) than for both jabberwocky (midline = .73 μV and lateral = −.08 μV, $p = .001$ and $p < .001$, respectively) and French (midline = .71 μV, $p < .001$, and lateral = −.26 μV, $p < .001$).

Table 3 – Results of separate ANOVAs for each congruity condition on mean amplitudes of the ERPs at midline and at lateral electrodes.

		Latency band (msec)	Factors	*df*	*F*	*p*
Strong incongruity	Midline	0–100	L	2, 56	3.75	*
		200–500	L × E	4, 112	27.18	***
		700–900	L × E	4, 112	5.64	**
	Lateral	0–100	L	2, 56	6.18	**
		200–500	L × R	4, 112	41.73	***
		700–900	L × R	4, 112	6.97	***
Weak incongruity	Midline	300–700	L × E	4, 112	4.32	*
	Lateral	200–300	L	2, 56	10.52	***
		300–400	L × H	2, 56	4.78	*
		300–700	L × R	4, 112	6.08	**
Congruous endings	Midline	300–700	L	2, 56	10.21	***
		300–700	L × E	4, 112	3.56	*
	Lateral	300–700	L	2, 56	12.01	***
		300–700	L × R	4, 112	4.99	**

L: language (Italian *vs* jabberwocky *vs* French); C: congruity (congruous *vs* weak incongruity *vs* strong incongruity); R: ROI (region of interest); H: hemisphere; E: electrodes.
df: Uncorrected degrees of freedom; *p*: probability level after Greenhouse–Geisser correction of *F*-values.
*: $p < .05$; **: $p < .01$; ***: $p < .001$.

For the weak Incongruity and at lateral electrodes, mean amplitudes in the 200–300 msec range were more negative (N2/N3 component) for Italian (−2.10 μV, $p < .001$) and jabberwocky (−1.68 μV, $p < .01$) than for French (−.45 μV; see Fig. 3). Moreover, the Language by Electrode and Language by ROI interactions were also significant in the 300–700 msec range. The amplitude of the positivity (P3b component) at parietal sites was larger to Italian (midline = 4.74 μV and lateral = 3.07 μV), than to jabberwocky (midline = 2.52 μV, $p < .001$, and lateral = 1.17 μV, $p < .001$) and to jabberwocky than to French (midline = 1.27 μV, $p < .05$ and lateral = .30 μV, $p < .05$). Finally, the Language by Hemisphere interaction was significant in the 300–400 msec latency band: the amplitude was significantly more negative at right (−2.56 μV) than at left (−1.34 μV, $p < .001$) electrodes only for jabberwocky.

For congruous endings and at both midline and lateral electrodes in the 300–700 msec latency band, both the main effect of Language and the Language by Electrodes and Language by ROI interactions were significant. The amplitude of the positivity (P3b component) at parietal sites was larger to Italian (midline = 3.69 μV and lateral = 2.40 μV) than to jabberwocky (midline = 2.16 μV, $p < .001$, and lateral = .99 μV, $p < .001$) and to jabberwocky than to French (midline = 1.32 μV and lateral = .27 μV, $p = .08$ and $p < .001$, respectively; see Fig. 4).

4. Discussion

Behavioral data showed a strong and reliable effect of language familiarity. Moreover, the level of performance for jabberwocky was higher than for French sentences, showing that participants were able to use the intonational contour of sentences to detect differences in pitch. Inspection of Fig. 1b shows that the strong incongruity condition is set apart from the other two conditions in the RTs data, suggesting that the congruity and weak incongruity conditions were more difficult to discriminate from each other. In particular, in the strong incongruity condition where the pitch change was more pronounced, participants were almost equally accurate in all three languages, but they were faster in Italian than in jabberwocky and French. Thus, although the pitch variation was strong, optimal performance was obtained when complete linguistic and semantic information was available.

As noted, the weakly incongruous condition was difficult to discriminate from the congruous condition, which is expected since participants were non-musicians (e.g., Schön et al., 2004). Although the congruous condition was almost as accurate as the strongly incongruous condition and both were at ceiling (Fig. 1a), the strongly incongruous condition was fast, while congruous and weakly incongruous conditions were slower (Fig. 1b). Even if context helped in setting up an expectation about the position of the last word, the latter conditions presumably required a re-analysis of the auditory input in order to decide if a pitch change had occurred. In this re-analysis, the pitch contour of the last word had to be compared to that of the sentence context. Thus, congruous and weakly incongruous conditions were most suitable to see effects of familiarity of the intonational patterns of the native language. In both conditions, accuracy was higher and RTs

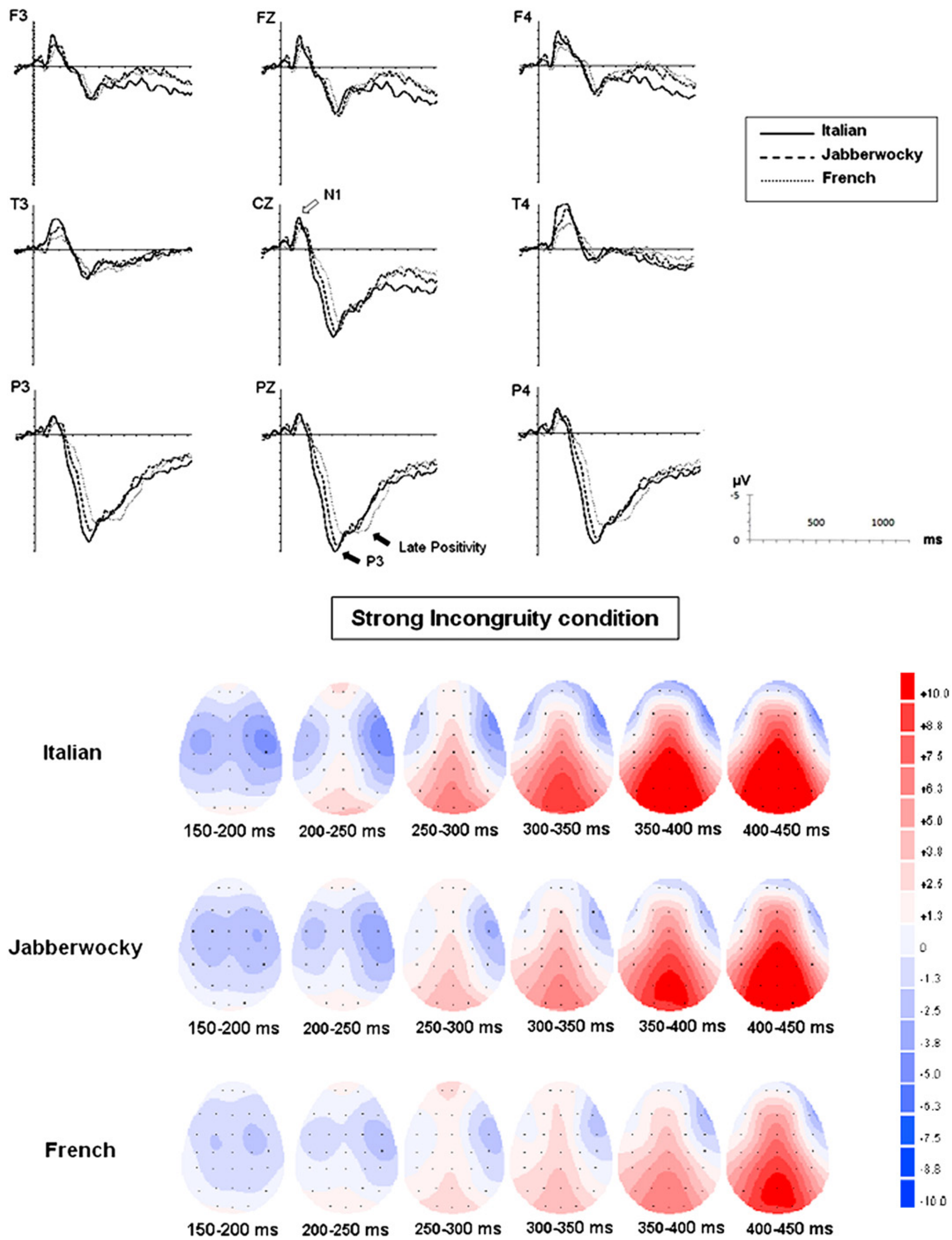

Fig. 2 – Upper panel: grand-averaged event-related potentials elicited by the strongly incongruous condition in Italian, jabberwocky and French sentences. Selected traces from 9 electrodes are presented. The N1 component is marked by a white arrow. The P3 component and late positivity are marked by black arrows. Lower panel: topographic maps of scalp electrical activity computed every 50 msec from 150 to 450 msec from the onset of the last word in the strongly incongruous condition for Italian (top), jabberwocky (middle) and French (bottom) sentences.

were faster for Italian and jabberwocky (that did not differ from one another) than in French. This pattern suggests that participants were able to exploit familiarity with the F0 contour of the sentences efficiently. Thus, the present results add to the evidence brought forward by Marques et al. through (a) a direct comparison of the maternal and foreign language in a within-subjects design, and (b) the inclusion of a pseudo-language (jabberwocky) in which only familiar prosodic information provides important clues to the pitch change. The former confirmed that a full semantic interpretation allows listeners to optimally perform the task of detecting a pitch change, even in conditions in which the pitch change is perceptually very strong. Although this finding may seem strange, it can be explained, as Marques et al. (2007) did, by assuming that a full semantic interpretation provides better clues to anticipate the sentence-final word. The latter finding

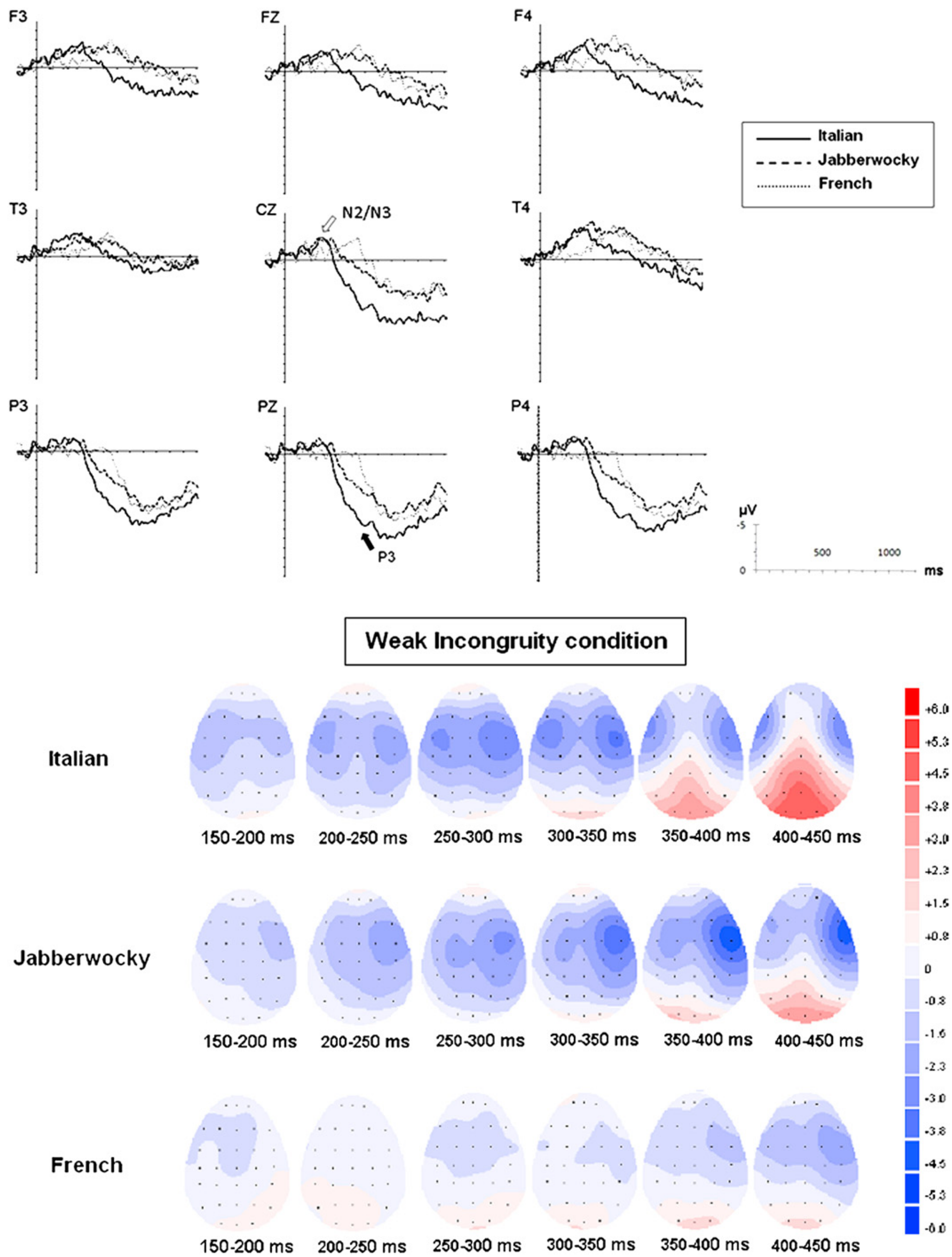

Fig. 3 – Upper panel: grand-averaged event-related potentials elicited by the weakly incongruous condition in Italian, jabberwocky and French sentences. Selected traces from 9 electrodes are presented. The N2/N3 component is marked by a white arrow. The P3 component is marked by a black arrow. Lower panel: topographic maps of scalp electrical activity computed every 50 msec from 150 to 450 msec from the onset of the last word in the weakly incongruous condition for Italian (top), jabberwocky (middle) and French (bottom) sentences.

shows that, although slightly slower, the jabberwocky condition allowed listeners to perform the task with optimal accuracy, and better than in a foreign language, based purely on familiarity of intonation.

The pattern of ERP data was in line with behavioral data. Overall, results showed that the amplitude of the positivity (P3 component) that developed at parietal sites 200/300 msec after final word onset was largest for Italian sentences, intermediate for jabberwocky sentences and smallest for French sentences. Thus, these analyses nicely complement behavioral ones, for which accuracy was at ceiling in the strongly incongruous condition. As noted above, the P3b component is taken to reflect post-perceptual processes involved in categorization and decision and P3b amplitude is larger when stimuli are

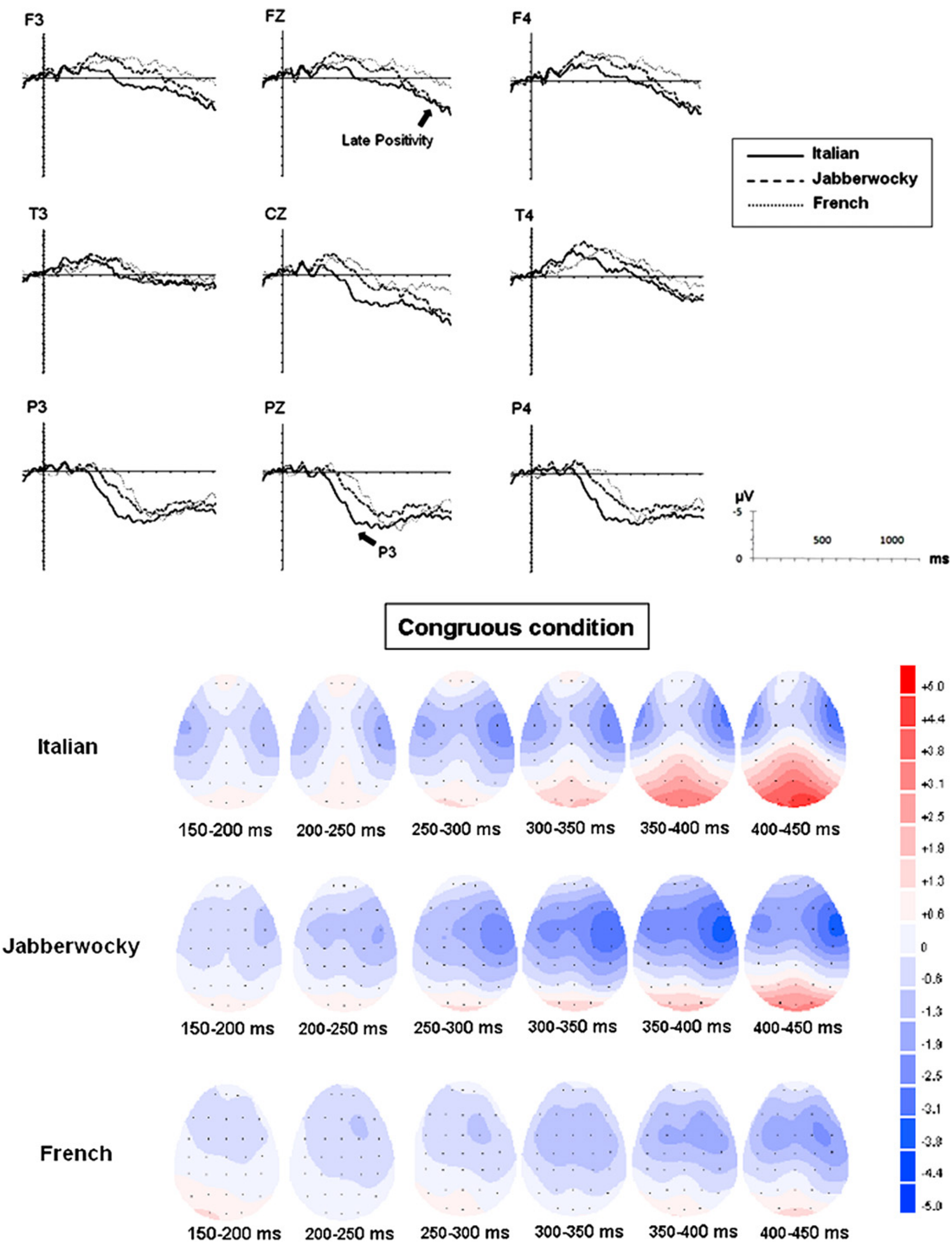

Fig. 4 – Upper panel: grand-averaged event-related potentials elicited by the congruous condition in Italian, jabberwocky and French sentences. Selected traces from 9 electrodes are presented. The P3 component and late positivity are marked by black arrows. Lower panel: topographic maps of scalp electrical activity computed every 50 msec from 150 to 450 msec from the onset of the last word in the congruous condition for Italian (top), jabberwocky (middle) and French (bottom) sentences.

easier to categorize (e.g., Hillyard et al., 1971). Final words were easier to categorize in Italian because of converging evidence from linguistic cues required for the semantic and syntactic interpretation of the sentence. In line with this interpretation, a late positivity developed for strong incongruities in French sentences in the 700–900 msec latency band over parietal sites, likely to reflect the greater difficulty of conducting such analysis in an unfamiliar language. The latency of this effect is similar to the onset latency of the late positivity (600–1000 msec) to strong incongruity in Portuguese sentences reported by Marques et al. (2007).

While effects, which appeared in the strong incongruity condition, might reflect automatic perception of a mismatch of the pitch variation with the preceding context (Okamoto et al., 2007; Sanders et al., 2009), it seems that congruous and weakly incongruous conditions needed more selective attention to the pitch differences. This interpretation is supported by the occurrence of an N2/N3, most visible in the most difficult condition, that is, the weakly incongruous condition in French (see Fujioka et al., 2006 and Moreno et al., 2009, for similar interpretations). A comparison of the maps of the three conditions (Figs. 2–4) supports this interpretation, with congruous and weakly incongruous conditions showing a similar pattern, while they both differ from the strongly incongruous condition.

In order to investigate the extent to which obtained effects were purely due to prosodic differences reflected in F0 variations, acoustic analyses were carried out on the materials. In particular, the F0 of the target word and of the two words before the last were measured in each language, in the congruous and weakly incongruous conditions. Preliminary analyses showed that in the congruous condition, in all three languages, there was a drop in F0 exactly at the onset of the last word and there was no significant difference in F0 values between the antepenultimate and penultimate word. In some sentences the penultimate word was an article, or a preposition, in some sentences an adjective or a verb. Thus the penultimate word did not always belong to the same syntactic clause as the target. Further analyses were carried out on the average F0 of the antepenultimate and penultimate words, combined. In the congruous condition the F0 values for Italian, jabberwocky and French, respectively, were 178, 175 and 194 Hz for the target word and 236, 239, and 239 Hz for the words before the target. The difference between preceding words and target was significant in all three languages. Specifically there was a 58 Hz difference for Italian [$t(98) = 24.04$, $SE = 2.42$, $p < .001$]; a 64 Hz difference for jabberwocky [$t(98) = 28.45$, $SE = 2.24$, $p < .001$]; and a 45 Hz difference for French [$t(98) = 19.84$, $SE = 2.28$, $p < .001$].

In the weakly incongruous condition, where a 115% increase in F0 was made, F0 of the target word increased to 204, 201, 223 Hz, respectively for Italian, jabberwocky and French. The difference between F0 of the target and that of preceding words was larger in Italian and jabberwocky than in French (32, 38, 16 Hz, respectively) but it was again significant in all three languages [$t(98) = 12.35$, $SE = 2.56$, $p < .001$; $t(98) = 15.84$, $SE = 2.38$, $p < .001$; $t(98) = 6.94$, $SE = 2.52$, $p < .001$].

These analyses showed that the trend of pitch was similar in all three languages, showing a significant drop in the last word, suggesting that acoustic information provided important cues for detecting prosodic variations. Moreover, the fact that this drop was larger in Italian and jabberwocky than in French may certainly explain part of the variance for the differences found in the behavioral analyses, where the performance of Italian and jabberwocky did not differ, while performance on French was lower. Further, these analyses show that acoustically Italian and jabberwocky did not differ. Thus a substantial part of the variance in the obtained effects was due to the acoustic difference in pitch between context words and target, reflecting prosodic variations. These variations were larger for Italian and jabberwocky than for French, and this added to another possible factor affecting the results, namely the familiarity with the pattern of acoustic variations in the native language.

A further issue is whether the different performance in the three language conditions might exclusively be due to differences in acoustic information in the two words before the target, rather than whole sentence processing. However, contrary to this claim, while the acoustic values of Italian and jabberwocky were very close, comparing the three language conditions, there emerged more or less strongly, but constantly, a pattern showing a linear trend for Italian, jabberwocky and French. This trend was apparent in both behavioral results (see Fig. 1) and ERP data (see Results section and Figs. 3 and 4). If performance of the participants had been based on the pitch variations of the target compared to only the preceding words, no difference between jabberwocky and Italian would be expected, as acoustically the two conditions did not differ. This difference was apparent, in contrast, in the ERP data. Considering both the congruent and weakly incongruent conditions, for example, the positive component was significantly less marked for jabberwocky than for Italian at parietal and frontal sites (see Results section and Figs. 3 and 4). Moreover, in the strong incongruity condition, the early N1 component was larger for Italian and jabberwocky sentences that did not differ acoustically, than for French sentences. This finding is interesting since the N1 component is typically taken to reflect sensory/perceptual analysis of the stimulus. Thus, perceptual processing of the strong incongruity seems to be influenced not only by the linguistic expectations derived from the sentence context (Italian sentences) but also by the familiarity with the pitch contour (jabberwocky sentences).

Similar to previous results by Schön et al. (2004), Mietz et al. (2008) also reported that both an early negative component, peaking around 120 msec after critical word onset, and a large positive component, starting around 500 msec after critical word onset (P600), developed following inadequate (and mismatching) prosody. The negative component was taken to reflect the detection of a mismatch between expectations created by the preceding prosodic context and the actual prosodic contour of the critical word, thereby suggesting the online influence of suprasegmental context. The positive component was taken to reflect processing of the mismatch and categorization processes. While the overall pattern of results was similar to that obtained in the present study, the critical word was always preceded by a pause, associated to specific syntactic structures. In the present study, in contrast, the pitch change matched the beginning of a lexical, not a prosodic word, and therefore never coincided with a syntactic/prosodic boundary. That is, the prosodic word comprising the final noun on which the pitch change was introduced also included at least an article, and most often adjectives or even verbs, and there was no pause, or silence, or change in intonation, before the onset of the final word, that might suggest that the critical pitch change would occur. Thus, the early negativity found in the present study, as well as in Schön et al. (2004) was not elicited by a syntactic/prosodic boundary, where the prosodic variation in pitch contour signals that a phrase ends and another one begins. The increased N1 amplitude reported here for Italian and jabberwocky sentences may rather reflect the segmentation of the auditory input in lexical words and/or increased attention to words with strong pitch changes (see also Sanders et al., 2009).

Several interpretations can be proposed regarding the functional significance of the increased N1 amplitude to strong incongruities at the end of Italian and jabberwocky sentences compared to French sentences: (a) processing of the acoustic differences triggered by pitch changes; (b) detection of a mismatch between the prosody of the sentence context and word onset on which a pitch change occurs; (c) segmentation of the last word from the context (Connolly et al., 1992; Connolly and Phillips, 1994; Sanders and Neville, 2002, 2003; Sanders et al., 2002); (d) selective attention processes associated to the pitch change and individuation of the critical word (Sanders et al., 2009). Obviously these hypotheses are not exclusive and might partially concur to the results.

In sum, the present study provided evidence that although a full (semantic, syntactic, prosodic) interpretation provides a better context for the detection of strong and weak F0 changes, familiarity with a prosodic contour provides strong support to perform the task. The similarity of results for Italian and jabberwocky sentences shows that participants were able to predict the upcoming pitch change from the prosodic contour of the sentence context, suggesting that familiarity with the prosodic contour is an important element in speech processing (Isel et al., 2005).

Acknowledgements

The authors wish to thank Daniele Schön for providing the French stimuli and the script for the pitch modification of the stimuli, Marianna Munafò for her assistance in ERP data analyses and Sandro Bettella for technical assistance.

REFERENCES

Astésano C, Besson M, and Alter K. Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research*, 18(2): 172–184, 2004.
Best C and Strange W. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, 20(3): 305–330, 1992.
Boersma P and Weenink D. Praat, a system for doing phonetics by computer, version 3.4. *Institute of Phonetic Sciences of the University of Amsterdam, Report* 132, 1996.
Cho T and McQueen JM. Phonological versus phonetic cues in native and non-native listening: Korean and Dutch listeners' perception of Dutch and English consonants. *The Journal of the Acoustical Society of America*, 119(5): 3085–3096, 2006.
Christophe A, Gout A, Peperkamp S, and Morgan J. Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics*, 31(3–4): 585–598, 2003.
Christophe A, Peperkamp S, Pallier C, Block E, and Mehler J. Phonological phrase boundaries constrain lexical access I. Adult data. *Journal of Memory and Language*, 51(4): 523–547, 2004.
Connolly JF, Phillips NA, Stewart SH, and Brake WG. Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences. *Brain and Language*, 43(1): 1–18, 1992.
Connolly JF and Phillips NA. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, 6 (3): 256–266, 1994.
Cutler A and Norris DG. The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14: 113–121, 1988.
Cutler A, Dahan D, and van Donselaar W. Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2): 141–201, 1997.
Di Cristo A and Hirst D. *Intonation Systems*. Cambridge: Cambridge University Press, 1998.
Dilley L and McAuley JD. Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language*, 59(3): 294–311, 2008.
D'Imperio M. Tonal alignment, scaling and slope in Italian question and statement tunes. *Travaux Interdisciplinaires Du Laboratoire Parole Et Langage d'Aix-En-Provence (TIPA)*, 21: 25–44, 2002.
D'Imperio M, Petrone C, and Nguyen N. Effects of tonal alignment on lexical identification in Italian. In Gussenhoven C and Riad T (Eds), *Tones and Tunes*. Berlin: Mouton de Gruyter, 2007: 79–106.
Eckstein K and Friederici AD. It's early: Event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *Journal of Cognitive Neuroscience*, 18(10): 1696–1711, 2006.
Flege JE. Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics*, 16(4): 425–442, 1995.
Friedrich CK, Alter K, and Kotz SA. An electrophysiological response to different pitch contours in words. *NeuroReport*, 12(15): 3189, 2001.
Fujioka T, Ross B, Kakigi R, Pantev C, and Trainor L. One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, 129: 2593–2608, 2006.
Hahne A and Friederici AD. Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, 11(2): 194–205, 1999.
Halliday MAK and Greaves W. *Intonation in the Grammar of English*. London; Oakville, CT: Equinox Pub., 2006.
Heim S and Alter K. Prosodic pitch accents in language comprehension and production: ERP data and acoustic analyses. *Acta Neurobiologiae Experimentalis*, 66(1): 55–68, 2006.
Hillyard SA, Squires KC, Bauer JW, and Lindsay PH. Evoked potential correlates of auditory signal detection. *Science*, 172: 1357–1360, 1971.
Isel F, Alter K, and Friederici AD. Influence of prosodic information on the processing of split particles: ERP evidence from spoken German. *Journal of Cognitive Neuroscience*, 17(1): 154–167, 2005.
Kuhl PK, Stevens E, Hayashi A, Deguchi T, Kiritani S, and Iverson P. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9(2): F13–F21, 2006.
Magne C, Astésano C, Lacheret-Dujour A, Morel M, Alter K, and Besson M. On-line processing of "pop-out" words in spoken French dialogues. *Journal of Cognitive Neuroscience*, 17(5): 740–756, 2005.
Magne C, Schön D, and Besson M. Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18(2): 199–211, 2006.
Magne C, Astésano C, Aramaki M, Ystad S, Kronland-Martinet R, and Besson M. Influence of syllabic lengthening on semantic processing in spoken French: Behavioural and electrophysiological evidence. *Cerebral Cortex*, 17(11): 2659–2668, 2007.
Marques C, Moreno S, Castro SL, and Besson M. Musicians detect pitch violation in a foreign language better than

nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, 19(9): 1453–1463, 2007.
Marslen-Wilson WS, Tyler LK, Warren P, Genier P, and Lee CS. Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 45A(1): 73–87, 1992.
Mietz A, Toepel U, Ischebeck A, and Alter K. Inadequate and infrequent are not alike: ERPs to deviant prosodic patterns in spoken sentence comprehension. *Brain and Language*, 104(2): 159–169, 2008.
Moreno S, Marques C, Santos A, Santos M, Castro SL, and Besson M. Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, 19: 712–723, 2009.
Morton J and Jassem W. Acoustic correlates of stress. *Language and Speech*, 8: 159–187, 1965.
Okamoto H, Stracke H, Wolters CH, Schmael F, and Pantev C. Attention improves population-level frequency tuning in human auditory cortex. *Journal of Neuroscience*, 27(39): 10383–10390, 2007.
Pannekamp A, Toepel U, Alter K, Hahne A, and Friederici AD. Prosody-driven sentence processing: An event-related brain potential study. *Journal of Cognitive Neuroscience*, 17(3): 407–421, 2005.
Rossi M. Intonation in Italian. In Hirst D and Di Cristo A (Eds), *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press, 1998: 219–238.
Sanders LD, Ameral V, and Sayles K. Event-related potentials index segmentation of nonsense sounds. *Neuropsychologia*, 47(4): 1183–1186, 2009.
Sanders LD and Neville HJ. Segmenting nonsense: An event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, 5(7): 700–703, 2002.
Sanders LD, Neville HJ, and Woldorff M. Speech segmentation of native and non-native speakers: The use of lexical, syntactic, and stress-pattern cues. *Journal of Speech, Language, and Hearing Research*, 45(3): 519–530, 2002.
Sanders LD and Neville HJ. An ERP study of continuous speech processing. I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, 15(3): 228–240, 2003.
Schafer AJ, Speer SR, Warren P, and White SD. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29(2): 169–182, 2000.
Schön D, Magne C, and Besson M. The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41(3): 341–349, 2004.
Steinhauer K, Alter K, and Friederici AD. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, 2(2): 191, 1999.
Steinhauer K and Friederici AD. Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, 30(3): 267–295, 2001.
Steinhauer K. Electrophysiological correlates of prosody and punctuation. *Brain and Language*, 86(1): 142–164, 2003.
Warren P, Grabe E, and Nolan F. Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes*, 10(5): 457–486, 1995.

Please cite this article in press as: Colombo L, et al., Detection of pitch violations depends upon the familiarity of intonational contour of sentences, Cortex (2010), doi:10.1016/j.cortex.2010.03.008