

Chapter 5

On the role of prosody in disambiguating *wh*-exclamatives and *wh*-interrogatives in Cosenza Italian

Olga Kellert

Georg-August-Universität Göttingen

Daniele Panizza

Georg-August-Universität Göttingen

Caterina Petrone

Laboratoire Parole et Langage, Aix-Marseille Université

This work investigates the role of prosody in the perception of *wh*-exclamatives and (information-seeking) *wh*-interrogatives in Cosenza Italian, a Southern Italian variety spoken in Calabria. Following recent research on prosody, we use a two-alternative forced-choice identification task in combination with reaction time measurements, as reaction times have been used as a better substitute of discrimination scores to investigate categorical perception of prosodic contrasts. Here, this methodology is preferred to more difficult offline tasks (e.g. the gating paradigm) to test to what extent phonetic/phonological cues distributed over the utterance might guide listeners' responses during sentence type identification. Our results show that listeners identify the two sentence types after the end of the utterance in most of the trials, not before it. This suggests that prosodic cues that occur before the end of the utterance (e.g. in the prenuclear section of the intonational contour) are not strong enough by themselves to guide the pragmatic interpretation of the utterances. Furthermore, our study shows that exclamatives are processed faster than interrogatives, but this effect disappears when segmental duration is taken into account.



1 Introduction

While in languages such as English, *wh*-exclamatives and information-seeking *wh*-interrogatives (hereafter “*wh*-interrogatives”) are syntactically differentiated (e.g. *How many books you have read!* vs. *How many books have you read?*), in Italian the two sentence types are syntactically the same:

(1) Italian *wh*-exclamatives

Quanti romanzi ha scritto la tua amica!
How many novels has written the your friend
‘How many novels your friend wrote!’

(2) Italian *wh*-interrogatives

Quanti romanzi ha scritto la tua amica?
How many novels has written the your friend
‘How many novels did your friend write?’

Studies on different languages have already shown that when syntactic structure is ambiguous, listeners mainly rely on prosodic information to identify exclamatives and interrogatives (cf. Batliner 1988; Eady & Cooper 1986; Soriano 2011; 2012; Gyuris et al. 2013).

However, it is still unclear to what extent temporally distributed phonological/phonetic properties are exploited by listeners for sentence-type identification. Consider the Italian examples in (1) and (2). The prosodic cues that determine the exclamative/interrogative meaning of these sentences could be contained in the *wh*-phrase (‘how many novels’), in the verb phrase (‘has written’), or in the final subject phrase (‘your friend’). In the variety spoken in Cosenza (Southern Italy), *wh*-exclamatives and *wh*-interrogatives contain different prosodic cues in both the prenuclear region (i.e. at the beginning of the intonation contour) and in the nuclear region (i.e., at the end of the intonation contour). *wh*-exclamatives exhibit a %H at the left edge of the intonational phrase, which is absent in *wh*-interrogatives (where a prenuclear accent H* is produced on the *wh*-constituent; see Soriano 2012 and §2 for details). Furthermore, the two sentence types are differentiated in the nuclear-accent choice associated with the verb phrase (L* in *wh*-exclamatives vs. L+H* in *wh*-interrogatives; *ibid.*).

Our study, focusing on Cosenza Italian, addresses two main questions. First, we ask whether listeners rely on the nuclear information alone or whether the prenuclear region also contributes to the perception of sentence type. Given that %H might differ from the prenuclear H* in many F0 dimensions (such as tonal

alignment and pitch excursion), we hypothesize that listeners should be able to discriminate the two sentence types from the beginning.

This question is linked to the issue in intonation research of how intonational meaning is created. The Autosegmental-Metrical (AM) framework assumes that the nuclear pitch accent is merely the last accent within a specific major prosodic phrase (cf. Ladd 2008 for a review). The overall meaning of a tune results from the independent contributions of its freely combinable, morpheme-like sub-parts, which include pitch accents and edge tones (Pierrehumbert & Hirschberg 1990). Despite this compositional-based approach to tune meaning, work within the AM theory often regards the nucleus – more or less implicitly – as the semantic “heart” of the tune. Hence, the current study searches for the meaning contribution of the prenuclear region and its potential interaction with the nuclear tune.

Furthermore, we ask whether the processing of the intonational contour is similar between the two sentence types or whether some cues are more salient in one sentence type than in the other. The %H is a marked pattern in Italian, since its use is restricted to a few contexts (including the *wh*-exclamatives). It also exhibits enhanced pitch excursion, which renders it perceptually salient (Sorianoello 2012). Hence, we hypothesize that the intonation of *wh*-exclamatives should be processed and identified much faster than the intonation of *wh*-interrogatives.

Concerning methodology, we employ an experimental design that involves an identification task with measurement of reaction times. Identification tasks have been already used in combination with reaction times (RTs) in prosody research, especially within the categorical perception paradigm (Chen 2003; Falé & Faria 2006; Niebuhr 2007; Feldhausen et al. 2011, among others). Reaction times are believed to reflect task difficulty (Massaro 1987) and they have been used as a substitute for discrimination scores to test whether the perception of prosodic contrasts is categorical or gradient (Chen 2003; Niebuhr 2007). This methodology has been applied to contrasts at the level of pitch accents (Chen 2003; Niebuhr 2007; Feldhausen et al. 2011), prosodic boundaries (Schneider 2011; Petrone et al. 2017), and global phonetic cues (such as pitch range; see Borràs-Comes et al. 2010). The studies in question are all based on manipulated stimuli. Given a continuum of manipulated stimuli, it is generally supposed that, in the case of categorical boundary perception, reaction times should be long at the location corresponding to the category boundary and short in other parts of the continuum. In contrast, a gradual increase in reaction time with proximity to the boundary might reflect a gradual increase of the ambiguity of the stimuli, pointing to a more gradient boundary perception. In the current study, we will measure identification scores and reaction times in response to natural (not manipulated) stimuli as a

first step in the investigation of an understudied prosodic contrast, i.e. the contrasts between *wh*-exclamatives and *wh*-interrogatives.

The structure of the paper is as follows: After a short review of the literature on *wh*-exclamatives and *wh*-interrogatives (§2), the goals of the study will be stated (§3) and a description of the experiment will be given (§4 and §5). The findings will be discussed in §6. Finally, the benefits and the challenges of the chosen methodology will be discussed in §7.

2 Previous investigations of the prosody of *wh*-exclamatives and *wh*-interrogatives

Prosodic characteristics of *wh*-interrogatives have been extensively studied in many languages (e.g. Zeng et al. 2004; Hedberg et al. 2010; Vanrell 2013). Concerning Italian, one of the most typical intonational characteristics of *wh*-interrogatives is the presence of utterance-final F0 fall (Chapallaz 1964; Avesani 1995; Soriano 2011). However, the intonational pattern of *wh*-interrogatives can vary in many respects depending on their pragmatic function (e.g. rhetorical vs. information-seeking; see Hedberg et al. 2010) and on the regional variety (see Avesani 1995; Soriano 2011; Gili Fivela et al. 2015, among others). In particular, in Cosenza Italian, the nuclear accent of information-seeking *wh*-interrogatives is associated with a rising accent L+H* followed by a low boundary tone L% (see Soriano 2011; Soriano et al. 2011).

Differently from *wh*-interrogatives, research on prosody of *wh*-exclamatives is cross-linguistically scarce (see Batliner 1988; Soriano 2011; Gyuris et al. 2013). Studies so far have rather focused on the syntactic and semantic differences between *wh*-exclamatives and *wh*-interrogatives (see Zanuttini & Portner 2003; Castroviejo 2006, among others). *wh*-exclamatives and *wh*-interrogatives would mainly differ with respect to factivity, i.e., exclamatives imply a true proposition and interrogatives do not. Zanuttini & Portner (2003) deduce the factivity of *wh*-exclamatives from two considerations. First, exclamatives (including *wh*-exclamatives) can be embedded under factive predicates only (Zanuttini & Portner 2003: 46):

- (3) Mary knows/*thinks/*wonders how very cute Mario is.

Furthermore, *wh*-exclamatives can never be used as questions and never induce a response from the interlocutor:

- (4) a. A: How tall is Mary? B: 1.80.
 b. A: How tall Mary is! B: #1.80.

Some studies have shown that prosodic information alone might be used to distinguish (*wh*- and non-*wh*) exclamatives from other sentence types. Exclamatives have been claimed to be cross-linguistically characterized by an initial extra-high pitch followed by a falling intonation contour (see O'Connor & Arnold 1961 for English; Delattre 1966 for French; D'Eugenio 1976 for Italian; Batliner 1988 for German). However, in some languages, such as Hungarian, exclamatives are instead characterized by an initial low pitch (a low boundary tone; see Gyuris et al. 2013). Concerning the nuclear region, Gyuris et al. (2013) found that Hungarian listeners classified acoustic stimuli as exclamatives and not as interrogatives if they included nuclear pitch accents with rising F0 pattern; delayed peaks; and a combination of low initial boundary tones and rising accents.

Concerning Cosenza Italian, Sorianello (2011) provides a phonological analysis of both *wh*-exclamatives and *wh*-interrogatives (see also Gili Fivela et al. (2015: 182) on *wh*-interrogatives in Cosenza Italian and other Italian varieties). The two sentence types are characterized by the same ending, i.e., a L-L% tonal sequence (see Figure 1 and Figure 2). However, they are differentiated phonologically in the regions preceding the utterance's final fall, i.e., both at the left edge of the intonational phrase (IP) and within the nuclear region. First, the initial part of *wh*-exclamatives is marked by a high boundary tone at the left edge of the IP (represented by %H in Figure 1), which is not present in *wh*-interrogatives (for %H in non-*wh*-exclamatives, see Sorianello 2012; Avesani 1995; Grice et al. 2005). *Wh*-interrogatives are characterized by a prenuclear accent L+ H* on the *wh*-constituent (see Figure 2). Moreover, the nuclear accent is specified differently, i.e., a low accent L* characterizes *wh*-exclamatives and a rising accent L+H* appears in *wh*-interrogatives. Along with the intonational cues, *wh*-exclamatives are also characterized by an extra lengthening of the duration of the nuclear-stressed syllable (Sorianello 2011).

While there are still no perception studies on *wh*-exclamatives in Cosenza Italian, there are studies of non-*wh*-exclamatives in other Southern varieties of Italian that suggest that the prenuclear region can contribute to their identification. In an experiment on Bari Italian, Sorianello (2012) used natural minimal pairs of non-*wh*-exclamatives and assertions, cutting the sentences at different temporal locations. This procedure is reminiscent of the gating procedure used in word-recognition research to investigate the uptake of acoustic-phonetic cues

to segmental structure (Grosjean 1980; Lahiri & Marslen-Wilson 1991, *inter alia*; see Petrone & D’Imperio 2011 for an example of this technique applied to international contrasts). Sorianello (2012) found that listeners could identify non-*wh*-exclamatives when only the initial part of the contour was available. By contrast, the prenuclear contour did not seem to provide robust cues for identifying assertions. Thus, the prenuclear cues are not equivalent across sentence types.

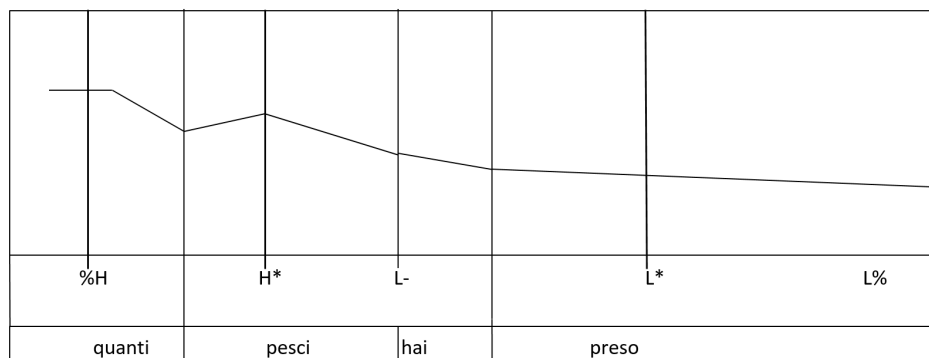


Figure 1: Our schematized version after Sorianello’s (2011: 316) representation of *Quanti pesci hai preso!* ‘How many fishes you took!’

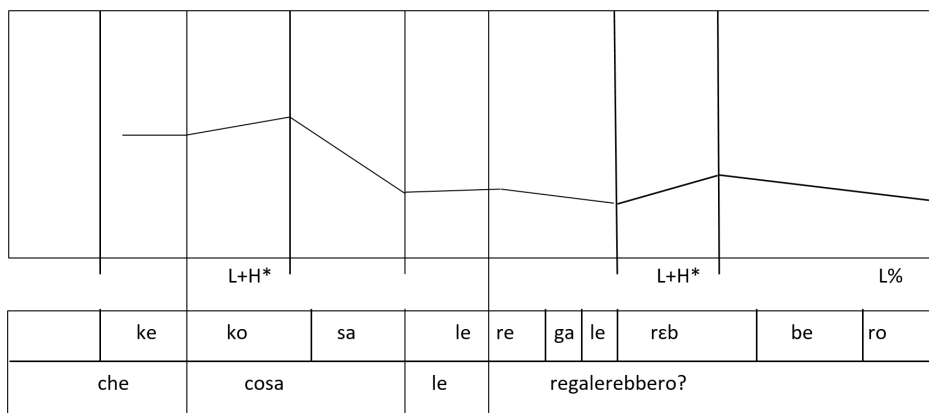


Figure 2: Our schematized version after the representation of Information-seeking *wh*-interrogative: *Che cosa le regalerebbero?* ‘What would they give her as a present?’ in (Gili Fivela et al. 2015: 182)

3 Goals of the study

The main goal of this study is to determine if prosodic cues play an important role in the disambiguation of *wh*-exclamatives and *wh*-interrogatives in Cosenza Italian. Given the phonetic/phonological contrasts between the two sentence types that were described in §2, listeners should be capable of distinguishing *wh*-interrogatives from *wh*-exclamatives on the basis of prosody (Hypothesis 1).

In particular, we seek to determine to what extent phonetic/phonological cues distributed over the utterance might guide listeners' responses. Given that the contrast between *wh*-exclamatives and *wh*-interrogatives should be noticeable already in the prenuclear region (Soriano 2011), we expect listeners to identify both sentence types before they hear the end of the sentence, with differences in the nuclear accent further helping perceptual disambiguation (Hypothesis 2).

Finally, we investigate the existence of potential differences in the processing of the two sentence types. Our hypothesis is that these processing differences do exist, since *wh*-exclamatives but not *wh*-interrogatives show a high boundary tone right at the beginning of the utterance, which is marked in Italian (i.e. %H at the left edge of the intonational phrase) (see Soriano 2011). We thus assume that the intonation of *wh*-exclamatives should be processed and identified much faster than the intonation of *wh*-interrogatives (Hypothesis 3).

In order to test the three hypotheses, we conducted a two-alternative forced-choice identification task combined with measurement of reaction times. Our expectation was that identification scores should depend on differences in prosody between the two sentence types. Specifically, cues in the prenuclear region should enable disambiguation between *wh*-exclamatives and *wh*-interrogatives, even though robust identification is only expected after the listener hears the entire utterance. Moreover, longer reaction times are expected to indicate higher uncertainty with respect to sentence-type identification, which should then be associated with lower identification scores in the identification task.

4 Experiment

4.1 Experimental stimuli

We constructed 20 pairs of morpho-syntactically and lexically identical *wh*-exclamatives and *wh*-interrogatives (like those in 1 and 2). Each target sentence had a syntactic structure as in (5): a complex *wh*-constituent (i.e., *quanti* 'how many' + noun); a verb phrase consisting of an auxiliary and a past participle; and a nominal constituent including a grammatical subject. (The definite article

preceding the subject constituent was omitted in the case of immediate family members, as typical in Italian.)

(5) Stimuli

[Quanti romanzi] [ha scritto] [la tua amica]

how many novels has written the your friend

'How many novels your friend wrote!' or

'How many novels did your friend write?'

In order to elicit different intonation patterns in *wh*-exclamatives and *wh*-interrogatives, we embedded the stimuli in pragmatic contexts that only matched one or the other sentence type, such as in (6) and (7):

- (6) *wh*-exclamative context: Your mother tells you that her friend spent 10 years of her life writing novels and shows you a list of her books. You exclaim:

Quanti romanzi ha scritto la tua amica!

how many novels has written the your friend

'How many novels your friend wrote!'

- (7) *wh*-interrogative context: Your mother tells you that her friend spent 10 years of her life writing novels. You ask your mom:

Quanti romanzi ha scritto la tua amica?

how many novels has written the your friend

'How many novels did your friend write?'

The sentences were produced by a 38-year-old female speaker from Cosenza. The speaker silently read the contexts and then uttered the sentences aloud. She was instructed to produce the sentences in a natural way.¹ No instructions were given as to what specific prosodic pattern to use in sentence production. The 40 target sentences were presented randomly and interspersed with 20 fillers: non-*wh*-interrogatives and non-*wh*-exclamatives such as 'Do you like coffee?' 'Open the door!' and 'I came late today' (a yes-no interrogative, an imperative, and a declarative, respectively). A complete list of the experimental target sentences and the fillers is given in the Appendix.

¹The discourse context was only presented in the production of the stimuli, to elicit a specific intonation contour, not in the identification task, since it would bias the responses of the listeners.

For the perception experiment, *wh*-exclamatives and *wh*-interrogatives constituted our auditory target stimuli. For each sentence, we identified four critical points or “marks” (M) at sequential temporal locations: M1 = the beginning of the utterance; M2 = the end of the *wh*-constituent; M3 = the end of the verb phrase; and M4 = the end of the subject constituent, which is also the end of the utterance. In example (8) below, M1 is represented by an opening bracket and M2–M4 by closing brackets.

- (8) [=M1 *Quanti romanzi*]=M2 *ha scritto*]=M3 *la tua amica*]=M4
 how many novels has written the your friend
 ‘How many novels your friend wrote!’ or
 ‘How many novels did your friend write!’

The four marks divided the experimental sentences into three regions. The initial region, between M1 and M2, consisted of the *wh*-constituent (e.g. *Quanti romanzi*); the middle region, between M2 and M3, consisted of the verb phrase (e.g. *ha scritto*); and the final region, from M3 to M4, consisted of the subject phrase at the end of the utterance (e.g. *la tua amica*).

4.2 Tasks and procedure

Since it has been shown that there are durational differences between the two sentence types (cf. Soriano 2011), something that was also clear from an initial analysis of our data, we measured the duration of each region. Statistical analysis of the duration showed that the initial region was significantly longer in the exclamative condition than in the interrogative one (on average, excl. = 616 ms vs. inter. = 546 ms, $p < .01$). The verb phrase was not significantly different between the two conditions (excl. = 592 ms vs. inter. = 604 ms, $p = .40$). The final subject phrase was longer in the exclamative condition (excl. = 933 ms vs. inter. = 751 ms, $p < .001$). Given that the duration of different regions seemed to play a crucial role in identification, this parameter was included as a variable in the statistical models (see §6).

Eighteen monolingual Italian native speakers (aged between 19 and 34 years) participated in the perception study and were reimbursed for their time. The group was composed of 10 women and 8 men. They were all from the Cosenza area and were either university students or employees at the *Università della Calabria*, in Cosenza. They reported no hearing problems.

Participants were asked to report which sentence type (*wh*-exclamative vs. *wh*-interrogative) better matched their auditory impression of the sentence. They

were instructed to carefully listen to the stimuli and to press a button as soon as they were certain of the sentence type.

This instruction was given in order to elicit as many early responses as possible, to enable us to check whether listeners could identify the sentence type before the end of the utterance (e.g. just from hearing the initial region, or the initial and middle regions). Before the experiment, listeners had a short practice session, with four practice sentences to identify. They did not receive any feedback on their answers.

The identification task lasted about 10 minutes for each listener. Their responses and reaction times (measured from the offset of the stimulus) were recorded. Stimulus presentation and response collection were performed by an open-source toolkit based on the Python module Pygame (cf. Peirce 2007 for an overview of PsychoPy, a toolkit based on the same system).²

Each trial began with a written question to the participant/listener asking if they were ready to start. A beep was used to signal that an utterance was about to start, in order to draw the participant's attention to the stimulus. The session began with the presentation of the auditory stimulus. For each listener, the identification task was broken into two blocks of 30 sentences, with some stimuli repeated in the second block. The goal of this manipulation was to check whether repetition of the same sentences influenced reaction times (Bentin & McCarthy 1994). Block 1 contained 10 *wh*-exclamative sentences and 10 *wh*-interrogative sentences that were not lexically identical ("non-minimal-pair condition"). Stimuli were presented in random order and interspersed with 10 fillers (non-*wh*-exclamatives and non-*wh*-interrogatives). Block 2 also contained 10 *wh*-exclamatives, 10 *wh*-interrogatives, and 10 fillers. This time, 5 of the 10 *wh*-exclamatives were lexically identical to 5 of the 10 *wh*-interrogatives presented in the first block, and likewise 5 of Block 2's *wh*-interrogatives were lexically identical to 5 of Block 1's *wh*-exclamatives. In other words, each participant heard 10 sentences under both the exclamative and the interrogative condition ("minimal-pair condition", e.g. 'How many novels has your friend written?' vs. 'How many novels your friend has written!'). In total, then, each participant heard 60 sentences: 20 *wh*-exclamatives, 20 *wh*-interrogatives, and 20 fillers. The stimuli were divided into four counter-balanced lists, to which listeners were randomly assigned. Listeners were tested one at a time in a quiet room.

²The module can be downloaded for free together with the data from this experiment (danielepanizza.org/pages/programming).

5 Statistical analysis and results

Before going into detail regarding the statistical analysis, we summarize our results with respect to the hypotheses we formulated in section 3. Our results show that:

Hypothesis 1 is confirmed. Listeners are very accurate in distinguishing *wh*-exclamatives from *wh*-interrogatives solely on the basis of prosody.

Hypothesis 2 is partially confirmed. Listeners can distinguish between these two sentence types before they hear the end of the sentence, but this pattern is very rare in our data; much more often, listeners identified the sentence type after the end of the sentence.

Hypothesis 3 is partially confirmed. Listeners are faster at identifying *wh*-exclamatives than *wh*-interrogatives. However, this effect is the result of durational differences between the two sentence types, so that the processing advantage for *wh*-exclamatives disappears when the segmental duration is taken into account.

5.1 Identification task

Accuracy of sentence-type identification was very high in both experimental conditions. Listeners correctly identified exclamatives in 93.4% of the trials and interrogatives in 93.7% of the trials. Although listeners were instructed to make their choice as soon as possible, the great majority of responses were provided after the end of the utterance (“late” responses: 90.7% for *wh*-exclamatives and 92.0% for *wh*-interrogatives). As shown in Table 1, in only 31 trials in the *wh*-exclamative condition and 27 trials in the *wh*-interrogative condition did listeners provide “early” responses (i.e., before the end of the utterance). Early responses were mostly correct, suggesting that some listeners are indeed able to discriminate the prosody of the two sentence types before the end of the utterance. For *wh*-exclamatives, the error rate for early responses was 19% compared to 5% for late responses, while for *wh*-interrogatives, the error rates were 4% for early responses and 6% for late responses. These results suggest that listeners were more prone to error when providing an early answer in response to *wh*-exclamatives than to *wh*-interrogatives. However, the difference between correct and incorrect early responses could not be assessed statistically because of the low number of observations.

We ran a generalized linear mixed model (GLMM) in order to analyze the accuracy of late responses. We adopted *sentence type* (interrogative vs. exclamative) and *block* (Block 1 vs. Block 2) as fixed factors and *item* (i.e. the lexical mate-

rial) and *participant* as random factors, with maximal random-effect structure (cf. Barr et al. 2013), that is, with the greatest possible number of free slopes and intercepts on both random factors, provided that the model converges. From this model, no significant difference in the accuracy of identification was revealed between *wh*-exclamative and *wh*-interrogative conditions ($\beta = 0.11$, $z = 0.32$, $p = .75$). The factor *block* was also not significant ($\beta = 0.26$, $z = 0.81$, $p = .42$), i.e., there was no effect of the repetition of the lexical material on the accuracy of the responses. A subsequent model was run with *sentence type* (interrogative vs. exclamative) and *response type* (early vs. late response) as fixed factors. The model confirmed no significant effects for *sentence type* for early responses (estimate = 1.65, $z = 1.24$, $p = .21$).

Table 1: Responses given in each time region.

	Exclamative condition		Interrogative condition	
	Correct	Incorrect	Correct	Incorrect
Initial region M1-M2	0	0	0	0
Middle region M2-M3	1	0	0	0
Final region M3-M4	24	6	26	1
After M4	289	16	290	20

5.2 Reaction Times (RTs)

We ran a statistical analysis on the RTs obtained from the identification task. Prior to the analysis, incorrect answers for both early and late responses were excluded and a logarithmic transformation was applied to the RTs to achieve a normal distribution (cf. Baayen 2008). The dependent variable was the RTs measured relative to the end of the sentence, which was positive for the trials in which listeners provided late responses and negative for the trials in which they provided early responses. After the logarithmic transformation, we excluded two outliers presenting a value that was greater than 3 standard deviations.

Statistical assessment was accomplished by applying a linear mixed model (LMM) to the RTs. We adopted *sentence type* as the main factor of interest and

item and *participant* as random factors, with maximal random-effects structure (cf. Barr et al. 2013). We also checked whether the factor *block* had any effect on RTs. The LMM showed a non-significant effect of *block*, i.e., repetition of sentences did not have any influence on RTs ($\beta = 0.03$, $t = -0.12$, $p = .91$). Furthermore, there was no interaction between *sentence type* and *block* ($\beta = 0.03$, $t = -0.08$, $p = .43$). This allowed us to drop the factor *block* from the remaining analyses. Instead, there was a difference across *sentence type*, with the *wh*-exclamatives being identified faster than the *wh*-interrogatives. In the exclamative condition, listeners took 525 ms on average from the end of the sentence to provide a correct response, whereas they took 639 ms in the interrogative condition. This difference is statistically significant ($\beta = 0.07$, $t = 3.43$, $p = .01$). In the next round of analyses, we checked whether RTs were different across the two sentence-type conditions, taking into account both early and late responses. When listeners provided an early response, they pressed the button on average 259 ms before the end of the sentence in the exclamative condition vs. 239 ms before the end of the sentence in the interrogative condition. When listeners answered after the end of the sentence, they took on average 604 ms in the exclamative condition vs. 718 ms in the interrogative condition. The statistical significance of these differences was assessed by running a LMM with *sentence type* (*wh*-exclamative vs. *wh*-interrogative) and *response type* (early vs. late responses) as fixed factors and *item* and *participant* as random factors, with random slopes and intercepts. We found significant differences involving both *sentence type* and *response type*. The effect of *response type* was significant, as expected ($\beta = 0.06$, $t = 3.52$, $p = .01$). The effect of *sentence type* was also significant for late responses ($\beta = 1.2$, $t = 19.42$, $p = .01$). However, the interaction between *sentence type* and *response type* was not significant ($\beta = 0.11$, $t = 0.83$, $p = .42$). This last result should be interpreted with caution, in that the number of observations involving early responses was very low (see Table 1). Figure 3 presents RTs across *sentence type*, in separate plots for early and late responses. As can be seen, RTs are different for late responses, while they are very similar for early responses.

Given that the duration of the stimuli was different across the two sentence types (see §4), we conducted another analysis that included the durations of the initial region (the *wh*-phrase), the middle region (the verb phrase), and the final region (the subject phrase) as covariates. This analysis addressed the question of whether the difference in RT between *wh*-exclamatives and *wh*-interrogatives revealed by the above analyses was actually caused by *sentence type* or whether it was an epiphenomenon resulting from the durational differences between the component regions of the *wh*-exclamatives and the *wh*-interrogatives. An LMM

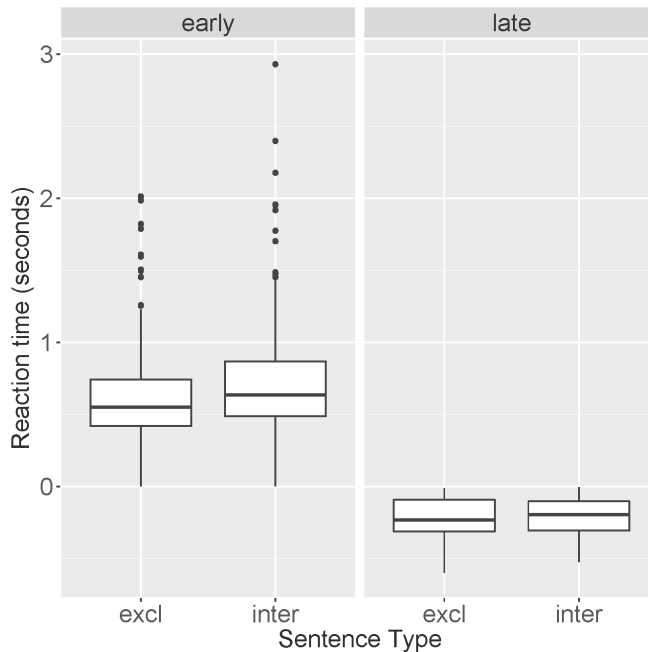


Figure 3: Reaction times (seconds) across *sentence type* and *response type*. The “0” value on the y-axis represents the end of the utterance, positive values represent RTs for responses given after the end of the utterance, and negative RTs correspond to responses given before the end of the utterance.

with *sentence type*, *response type*, and duration of *initial region*, *middle region*, and *final region* as fixed factors and *item* and *participant* as random factors yields the following results. There is no significant difference across *sentence type* ($\beta = 0.01$, $t = -0.77$, $p = .44$) nor is there any interaction between *sentence type* and *response type* ($\beta = 0.1$, $t = -1.08$, $p = .28$). Instead, *response type* ($\beta = 1.2$, $t = 21.9$, $p < .001$), *initial region* ($\beta = -0.31$, $t = -4.15$, $p < .001$), *middle region* ($\beta = -0.23$, $t = -2.56$, $p < .02$) and *final region* ($\beta = -0.33$, $t = -5.51$, $p < .001$) are significant. Hence, the results of this analysis show that the duration of each of the three regions of the sentence is a significant predictor of *reaction times*, while the main factor of our experimental design, *sentence type*, is not significant.

6 Discussion and conclusions

The identification task has shown that (Cosenza) Italian listeners are capable of distinguishing between *wh*-exclamatives and *wh*-interrogatives on the basis of

prosody. The fact that listeners gave correct responses for both *wh*-exclamatives and *wh*-interrogatives in more than 90% of trials indicates that there must be some prosodic marker that guides listeners' judgments. Furthermore, this experiment was a preliminary attempt to find out whether prenuclear cues (like the %H vs H* difference at the left edge of the IP) might be used for the purpose of pragmatic interpretation. The fact that our experiment elicited some early responses, roughly to the same degree in *wh*-exclamatives (8.3%) and *wh*-interrogatives (9%), is compatible with the hypothesis that listeners can employ prosodic information either in the nuclear region or in the prenuclear region to identify the sentence type. However, our results strongly support the hypothesis that the most relevant phonetic/phonological cues for sentence-type disambiguation are located at the end of the utterance, given that a) listeners gave their responses mostly (in more than 90% of the cases) after the end of utterance; b) the early responses were provided on average about 200 ms before the end of the utterance and more than 2 seconds after the offset of the region containing the nuclear cues; and c) phonetic/phonological cues of the final region significantly affected the RTs.

However, the fact that the duration of the initial and middle regions also significantly affected RT is strongly indicative that prosodic information in the prenuclear section might be exploited by listeners in identifying sentence type. If we take into account this last result, we might interpret the high rate of late responses as a result of listeners' insecurity about their decision. Given that they were instructed to be both fast and accurate, listeners may have collected phonetic/phonological cues while listening to the utterance in order to increase the probability of a reliable response.³

To conclude, our study has yielded important preliminary results concerning the identification and processing of *wh*-exclamatives and *wh*-interrogatives on the basis of prosody in Cosenza Italian. Listeners wait until the end of the utterance to respond in most but not all of the trials, thus suggesting that prosodic cues in the prenuclear contour are not strong enough by themselves to guide the pragmatic interpretation of the utterances in an unambiguous way. Furthermore, our study indicates that *wh*-exclamatives are processed faster than *wh*-interrogatives, but this effect disappears when the duration of different segmental regions of the utterance is taken into account.

³We thank a reviewer for pointing out this hypothesis to us.

7 Benefits and challenges of the chosen method

By combining an identification task with reaction time measurements, we measured not only accuracy in prosodic disambiguation (through the identification scores) but we also provided additional information about when exactly sentence-type identification takes place and whether one sentence type is more difficult to process than another (through reaction times).

We preferred this method to other offline tasks like the gating paradigm. In a gating paradigm, fragments of speech are presented to listeners in an order of increasing duration. Listeners usually have to identify these fragments and to rate their level of confidence. This method allows to obtain the location of the phonetic or phonological cue that is responsible for the correct identification (Prieto 2012). The gating paradigm has already been used in prosody research, particularly for investigating the contribution of the prenuclear contour to tune meaning (e.g. Petrone 2008; Petrone & D'Imperio 2011; Soriano 2012; Prieto 2012). One potential concern about this task is that stimuli consist in short pieces of artificially cut utterances, which might sound unnatural to listeners. Hence, the identification task might be more difficult to be accomplished. Furthermore, such a paradigm does not allow to track the continuous time course of utterance interpretation.

In the current study, we tried to overcome these concerns by using uncut stimuli, for which we tracked the course of utterance interpretation by means of reaction times. This seemed to us a simple technique that might be well suited for preliminary investigation of prosodic contrasts, such as the prosodic differences between *wh*-exclamatives and *wh*-interrogatives. However, our method did not clearly say whether the relevant phonological cues for sentence-type identification are included in the prenuclear region of the utterance or in the nuclear region. A potential problem with measuring RTs in an identification task in which listeners have to press a button is that there is a time delay between sentence-type identification and the response reaction (i.e., pressing the button). To address this issue, alternative methods for registering the response reaction could be implemented.

Online methods like eye-tracking or mouse tracking could be used (Marslen-Wilson et al. 1992; Pynte & Prieur 1996; Tomlinson & Fox Tree 2011; Warren 2014). In particular, a methodological challenge concerning eye-tracking would be to develop a paradigm to investigate the time course of processing of intonational meaning (i.e., taking into account the meaning contribution of the prenuclear and nuclear sections) during visual search. For instance, in a study on American

English, Heeren et al. (2015) created a “targeted language game” by which an indirect association between different sentence types (questions and statements) and referents (objects on a visual display) is created. Analysis of gazes demonstrated that listeners can make an immediate use of the nuclear accents and boundary tones in processing questioning vs. asserting utterances. Their stimuli were however limited to less typical syntactic constructions of American English (elliptical utterances only containing the nucleus), so the contribution of the prenuclear section was unclear. In an eye-tracking experiment on French, Petrone et al. (2016) combined long utterances produced with different degrees of commitment (as signaled by prosody: yes/no questions vs. incredulity questions) with pictures showing the corresponding facial expressions. Results indicated that French listeners can make immediate use of prenuclear cues for processing speaker commitment and that the effect of nuclear and prenuclear cues may vary across different utterance types. This kind of results encourages us to use the visual world paradigm to assess the use of intonation during sentence modality processing.

With regard to stimuli selection, the current experiment is based on natural stimuli and does not allow us to distinguish which acoustic marker contributed to listeners’ relatively high performance on sentence-type disambiguation at the end of the utterance. Our speaker produced *wh*-exclamatives and *wh*-interrogatives based on specific pragmatic contexts, but she was not asked to produce a specific set of intonation contours.

In future studies, we will investigate the influence of intonational cues (edge tones and pitch accents) by controlling for the tonal structure of the target sentences. When looking at intonational cues, durational differences could be controlled for by using resynthesized stimuli with similar segmental duration for *wh*-exclamatives and *wh*-interrogatives. Furthermore, a continuum of resynthesized stimuli could be used in order to determine whether the prosodic parameters under investigation are perceived in a categorical or gradient manner (see Niebuhr 2007).

Appendix: Data Set

Table 2: Itemlist_Fillers

Vieni stasera?	‘Will you come tonight?’
Mi fai un caffè?	‘Would you prepare me a coffee?’
Apriresti la finestra?	‘Would you open the window?’
Piove tanto?	‘Does it rain a lot?’
Sei bella?	‘Are you beautiful?’
Mi daresti il tuo numero?	‘Would you tell me your number?’
Perchè piangi?	‘Why do you cry?’
Hai visto il mio ragazzo?	‘Did you see my boyfriend?’
Hai 25 anni?	‘Are you 25 years old?’
Sei una stronza!	‘You are stupid?’
C’è qualcuno al telefono!	‘There is someone on the phone!’
Forse hai ragione!	‘Maybe you are right!’
Sei una persona speciale!	‘You are a special person!’
Vieni sta sera!	‘Come tonight!’
Guarda ’sto video!	‘Watch this video!’
Sei bellissima!	‘You are beautiful!’

5 *Wh-exclamatives and wh-interrogatives in Cosenza Italian*

Table 3: Duration markers according to target sentences (Interrogatives)

Target sentences	Placement of dur. markers (sec)			
	M1	M2	M3	M4
Quanti romanzi ha scritto la tua amica 'How many novels did your friend write?'	1.052	1.733	2.252	3.127
Quanti libri ha pubblicato il tuo professore 'How many books did your professor publish?'	1.367	1.828	2.507	3.596
Quante sigarette ha fumato papa 'How many cigarettes did Dad smoke?'	1.219	1.926	2.595	3.074
Quanti paesi ha visto tua sorella 'How many countries did your sister see?'	1.141	1.651	2.138	2.847
Quante cose ha aggiustato tuo padre 'How many things did your father adjust?'	1.421	1.996	2.587	3.169
Quante birre ha bevuto la tua amica 'How many beers did your friend drink?'	0.999	1.490	2.052	2.747
Quanti chili ha perso tuo nipote 'How many kilos did your nephew loose?'	0.895	1.351	1.870	2.613
Quanti corsi ha seguito tua sorella 'How many lectures did your sister take?'	1.094	1.577	2.222	3.069
Quanta torta ha mangiato tua sorella 'How much cake did your sister eat?'	1.090	1.618	2.222	3.059
Quanti libri ha comprato tuo padre 'How many books did your father buy?'	0.809	1.285	1.965	2.710
Quanti soldi ti ha dato tuo padre 'How much money did your father give you?'	1.186	1.778	2.280	2.935
Quanti vestiti ha disegnato il tuo amico 'How many clothes did your friend design?'	1.049	1.684	2.285	3.100
Quanti pesci ha pescato tuo fratello 'How much fish did your brother catch?'	1.130	1.626	2.249	3.113
Quanti cd ha inciso tuo zio 'How many discs did your uncle record?'	1.124	1.739	2.256	3.013
Quante arance ha raccolto tuo nonno 'How many oranges did your grandfather collect?'	1.335	1.860	2.494	3.169
Quanti quadri ha dipinto tua zia 'How many paintings did your aunt paint?'	1.197	1.688	2.325	3.055
Quanti dolci ha preparato tua madre 'How many sweets did your mother prepare?'	1.230	1.751	2.402	3.149
Quanti fiori ha piantato tua nonna 'How many flowers did your grandmother plant?'	1.413	1.887	2.492	3.255
Quante farfalle ha catturato tuo fratello 'How many butterflies did your brother capture?'	0.624	1.266	1.926	2.759
Quante scarpe ha comprato tua zia 'How many shoes did your aunt buy?'	1.219	1.780	2.476	3.063

Table 4: Duration markers according to target sentences (Exclamatives)

Target sentences	Placement of dur. markers (sec)			
Quanti romanzi ha scritto la tua amica 'How many novels your friend wrote!'	1.286	1.985	2.575	3.438
Quanti libri ha pubblicato il tuo professore 'How many books your professor published!'	1.087	1.655	2.356	3.570
Quante sigarette ha fumato papa 'How many cigarettes Dad smoked!'	1.098	1.960	2.737	3.348
Quanti paesi ha visto tua sorella 'How many countries your sister saw!'	1.076	1.600	2.062	2.970
Quante cose ha aggiustato tuo padre 'How many things your father adjusted!'	1.030	1.553	2.235	3.044
Quante birre ha bevuto la tua amica 'How many beers your friend drunk!'	1.097	1.619	2.174	3.180
Quanti chili ha perso tuo nipote 'How many kilos your nephew lost!'	0.891	1.307	1.800	2.673
Quanti corsi ha seguito tua sorella 'How many lectures your sister took!'	1.207	1.766	2.237	3.237
Quanta torta ha mangiato tua sorella 'How much cake your sister ate!'	1.068	1.645	2.158	3.032
Quanti libri ha comprato tuo padre 'How many books your father bought!'	0.686	1.420	2.068	3.147
Quanti soldi ti ha dato tuo padre 'How much money your father gave you!'	1.037	1.803	2.265	3.062
Quanti vestiti ha disegnato il tuo amico 'How many clothes your friend designed!'	0.942	1.828	2.454	3.361
Quanti pesci ha pescato tuo fratello 'How much fish your brother catch!'	2.031	2.638	3.213	4.245
Quanti cd ha inciso tuo zio 'How many discs your uncle recorded!'	1.143	1.886	2.346	3.226
Quante arance ha raccolto tuo nonno 'How many oranges your grandfather collected!'	1.750	2.301	2.884	3.789
Quanti quadri ha dipinto tua zia 'How many paintings your aunt painted!'	1.711	2.214	2.828	3.753
Quanti dolci ha preparato tua madre 'How many sweets your mother prepared!'	1.482	2.028	2.671	3.680
Quanti fiori ha piantato tua nonna 'How many flowers your grandmother planted!'	1.335	1.802	2.463	3.486
Quante farfalle ha catturato tuo fratello 'How many butterflies your brother captured!'	0.803	1.457	2.129	3.141
Quante scarpe ha comprato tua zia 'How many shoes your aunt bought!'	1.054	1.670	2.316	3.240

References

- Avesani, Cinzia. 1995. ToBI. Un sistema di trascrizione per l'intonazione italiana. In Gianni Lazzari (ed.), *Metodologie di analisi e di descrizione delle caratteristiche prosodiche e intonative dell'italiano. Atti delle V Giornate di Studio del Gruppo di Fonetica Sperimentale (A.I.A.)* Vol. XXII, 85–98. Povo (TN), Italy.
- Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Barr, Dale, Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory Language* 68. 255–278.
- Batliner, Anton. 1988. Produktion und Prädikation. Die Rolle intonatorischer und anderer Merkmale bei der Bestimmung des Satzmodus. In Hans Altmann (ed.), *Intonationsforschungen*, 207–221. Tübingen: Niemeyer.
- Bentin, Shlomo & Gregory McCarthy. 1994. The effects of immediate stimulus repetition on reaction time and event-related potentials in tasks of different complexity. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 20(1). 130–149.
- Borràs-Comes, Joan, Maria del Mar Vanrell & Pilar Prieto. 2010. The role of pitch range in establishing intonational contrasts in Catalan. In Mark Hasegawa-Johnson et al. (ed.), *Proceedings of the 5th International Conference on Speech Prosody*. Chicago, IL.
- Castroviejo, Elena. 2006. *Wh-Exclamatives in Catalan*. Universitat de Barcelona PhD.
- Chapallaz, Marguerite. 1964. Notes on the Intonation of Questions in Italian. In D. Abercrombie, D. B. Fry, P. A. D. MacCarthy, N. C. Scott & J. L. M. Trim (eds.), *In Honor of Daniel Jones*, 306–312. Longman.
- Chen, Aoju. 2003. Reaction time as an indicator to discrete intonational contrasts in English. In *Proceedings of Eurospeech 2003*, 97–100. Geneva.
- D'Eugenio, Antonio. 1976. The intonation systems of Italian and English. *Rassegna Italiana di Linguistica Applicata* 8(1). 57–85.
- Delattre, Pierre. 1966. Les 10 intonations de base du français. *The French Review* 40(1). 1–14.
- Eady, Stephen J. & William E. Cooper. 1986. Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America* 80. 402–415.
- Falé, Isabel & Isabel Hub Faria. 2006. Categorical perception of intonational contrasts in European Portuguese. In Rüdiger Hoffmann & Hansjörg Mixdorff

- (eds.), *Proceedings of the 3rd International Conference on Speech Prosody*, 69–72. Dresden.
- Feldhausen, Ingo, Andrea Pešková, Elena Kireva & Christoph Gabriel. 2011. Categorical perception of Porteño nuclear accents. In *Proceedings of the 17th International Congress of Phonetic Sciences*, 17–21. Hong Kong.
- Gili Fivela, Barbara, Cinzia Avesani, Marco Barone, Giuliano Bocci, Claudia Crocco, Mariapaola D'Imperio, Rosella Giordano, Giovanna Marotta, Michelina Savino & Patrizia Sorianello. 2015. Varieties of Italian and their intonational phonology. In Sónia Frota & Pilar Prieto (eds.), *Intonation in Romance*, 140–197. Oxford University Press.
- Grice, Martine, Mariapaola D'Imperio, Michelina Savino & Cinzia Avesani. 2005. Strategies for Intonation Labelling across Varieties of Italian. In Sun-Ah Jun (ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing*, 362–389. Oxford: Oxford University Press.
- Grosjean, François. 1980. Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics* 28. 267–283.
- Gyuris, Beáta, Katalin Mády & Ádám Szalontai. 2013. *Experimental investigations on the prosody of Hungarian exclamatives*. Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, 30 May, 2013.
- Hedberg, Nancy, Juan M. Sosa, Emrah Görgülü & Morgan Mameni. 2010. The prosody and meaning of wh-questions in American English. In Mark Hasegawa-Johnson et al. (ed.), *Proceedings of the 5th International Conference on Speech Prosody*. Chicago, Illinois.
- Heeren, Willemijn F. L., Sarah A. Bibyk, Christine Gunlogson & Michael K. Tanenhaus. 2015. Asking or Telling—Real-time Processing of Prosodically Distinguished Questions and Statements. *Language and Speech* 58(4). 474–501. DOI:10.1177/0023830914564452
- Ladd, D. Robert. 2008. *Intonational phonology*. 2nd edition. Cambridge: Cambridge University Press.
- Lahiri, Aditi & William Marslen-Wilson. 1991. The mental representation of lexical form. A phonological approach to the recognition lexicon. *Cognition* 38. 245–294.
- Marslen-Wilson, William, Lorraine K. Tyler, Paul Warren, Paula Grenier & Catherine S. Lee. 1992. Prosodic effects in minimal attachment. *Quarterly Journal of Experimental Psychology* 45. 73–87.
- Massaro, Dominic W. 1987. *Speech perception by ear and eye. A paradigm for psychological inquiry*. Mahwah, NJ, USA: Erlbaum.

- Niebuhr, Oliver. 2007. Categorical perception in intonation. A matter of signal dynamics? In *Proceedings of the 8th Annual Conference of the International Speech Communication Association 2007 (INTERSPEECH 2007)*, 109–112. Antwerp, Belgium: Curran Associates, Inc.
- O'Connor, Joseph D. & Gordon Frederick Arnold. 1961. *Intonation of colloquial English: A practical handbook*. 1. publ. London: Longman.
- Petrone, Caterina. 2008. *From targets to tunes. Nuclear and prenuclear contribution in the identification of intonation contours in Italian*. Laboratoire Parole et Langage, Université de Provence PhD.
- Petrone, Caterina & Mariapaola D'Imperio. 2011. From Tones to Tunes: Effects of the f0 Prenuclear Region in the Perception of Neapolitan Statements and Questions. In Sónia Frota, Gorka Elordieta & Pilar Prieto (eds.), *Prosodic Categories: Production, Perception and Comprehension* (Studies in natural language and linguistic theory), 207–230. Dordrecht: Springer Netherlands.
- Petrone, Caterina, Alessandra Lo Nobile, Christelle Zielinsky & Kiwako Ito. 2016. The role of prosody in processing speaker commitment in French. In Jonathan Barnes, Alejna Brugos, Stefanie Shattuck-Hufnagel & Nanette Veilleux (eds.), *Proceedings of the 8th International Conference on Speech Prosody*. Boston, USA. <https://hal.archives-ouvertes.fr/hal-01462229>.
- Petrone, Caterina, Hubert Truckenbrodt, Caroline Wellmann, Julia Holzgrefe-Lang, Isabell Wartenburger & Barbara Höhle. 2017. Prosodic boundary cues in German. Evidence from the production and perception of bracketed lists. *Journal of Phonetics* 61. 71–92.
- Pierrehumbert, Janet B. & Julia B. Hirschberg. 1990. The Meaning of Intonational contours in the Interpretation of Discourse. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in Communication*, 271–311. Cambridge MA: MIT Press.
- Prieto, Pilar. 2012. Experimental methods and paradigms for prosodic analysis. In Abigail C. Cohn, Cécile Fougeron & Marie K. Huffman (eds.), *The Oxford Handbook of Laboratory Phonology* (Oxford Handbooks in Linguistics), 528–538. Oxford: Oxford University Press.
- Pynte, Joel & Benedicte Prieur. 1996. Prosodic breaks and attachment decisions in sentence parsing. *Language and Cognitive Processes* 11. 165–191.
- Schneider, Edgar W. 2011. *English around the world. An introduction*. Cambridge: Cambridge University Press.
- Sorianello, Patrizia. 2011. Aspetti prosodici e pragmatici dell'atto esclamativo. *Studi Linguistici e Filologici Online* 9. 287–332.

- Sorianello, Patrizia. 2012. A prosodic account of Italian exclamative sentences. A gating test. In Qi-uwu Ma, Hongwei Ding & Daniel Hirst (eds.), *Proceedings of the 6th International Conference on Speech Prosody*, vol. 1, 298–301. Shanghai, China: Tongji University Press.
- Sorianello, Patrizia, Riccardo Giordano & Caterina Petrone. 2011. *L'Italiano parlato a Cosenza*. PPT presented at the workshop ToBI Italiano CRIL, Lecce.
- Tomlinson, John M. & Jean E. Fox Tree. 2011. Listeners' comprehension of uptalk in spontaneous speech. *Cognition* 119(1). 58–69.
- Vanrell, Maria del Mar. 2013. Pitch accent types and the perception of focus in Majorcan Catalan wh-questions. In Sylvie Hancil & Daniel Hirst (eds.), *Prosody and Iconicity*, vol. 13 (Iconicity in Language and Literature), 127–148. Amsterdam: John Benjamins Publishing Company. DOI:10.1075/ill.13.07van
- Warren, Paul. 2014. Sociophonetic and prosodic influences on judgements of sentence type. In Jennifer Hay & Emma Parnell (eds.), *Proceedings of the 15th Australasian International Conference on Speech Science and Technology*, 185–188. Christchurch.
- Zanutini, Rafaela & Paul Portner. 2003. Exclamative Clauses at the Syntax-Semantics Interface. *Language* 79(1). 39–81.
- Zeng, Xiao-Li, Philippe Martin & Georges Boulakia. 2004. Tones and intonation in declarative and interrogative sentences in Mandarin. In *Proceedings of the International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages, Beijing, China, March 28-31*, 235–238.